

Parameter and state estimation of nonlinear systems using a multi-observer under the supervisory framework

Michelle S. Chong, Dragan Nešić, Romain Postoyan and Levin Kuhlmann

Abstract

We present a hybrid scheme for the parameter and state estimation of nonlinear continuous-time systems, which is inspired by the supervisory setup used for control. State observers are synthesized for some nominal parameter values and a criterion is designed to select one of these observers at any given time instant, which provides state and parameter estimates. Assuming that a persistency of excitation condition holds, the convergence of the parameter and state estimation errors to zero is ensured up to a margin, which can be made as small as desired by increasing the number of observers. To reduce the potential computational complexity of the scheme, we explain how the sampling of the parameter set can be dynamically updated using a zoom-in procedure. This strategy typically requires a fewer number of observers for a given estimation error margin compared to the static sampling policy. The results are shown to be applicable to linear systems and to a class of nonlinear systems. We illustrate the applicability of the approach by estimating the synaptic gains and the mean membrane potentials of a neural mass model.

I. INTRODUCTION

The estimation of states and parameters is a long-standing problem in control theory, which is particularly involved when dealing with nonlinear systems. While approaches are available to estimate the states (*e.g.* [5], [6], [9], [23], [25]) or the parameters (*e.g.* [1], [19, Chapter 4], [29]), the simultaneous estimation of parameter and state remains a challenging problem [9]. A classical technique consists of augmenting the state vector with the parameter vector

M. Chong, D. Nešić and L. Kuhlmann are with the Department of Electrical and Electronic Engineering, the University of Melbourne, Australia. {chongms,dnesic,levink}@unimelb.edu.au

R. Postoyan is with the Université de Lorraine, CRAN, UMR 7039 and the CNRS, CRAN, UMR 7039, France. He is financially supported by the ANR under the grant SEPICOT (ANR 12 JS03 004 01). romain.postoyan@univ-lorraine.fr

and thus reducing the problem to state estimation only. However, this complicates the problem even for linear systems since rewriting the system in these new coordinates can turn the system highly nonlinear. We note that while nonlinear Kalman filters are often used in practice, their convergence is only provable under limited conditions [4, Chapter 8], [21], in addition, these filters may be difficult to tune.

In this paper, we investigate the problem of state and parameter estimation of deterministic nonlinear continuous-time systems by adopting an architecture known as the supervisory framework, see [28, Chapter 6]. We assume that the system parameters are constant and that they belong to a known, compact set. We sample the parameter set to form a finite set of nominal values. A state observer is designed for each of these nominal values, which is assumed to satisfy a given robustness property with respect to parameter mismatches. They form the multi-observer unit [28, Chapter 6]. The supervisory unit then provides state and parameter estimates by selecting one observer from the bank at any given time instant. We call this setup a *supervisory observer*. The parameter and state estimates are guaranteed to converge to the true values up to any given margin, provided that the number of observers is sufficiently large. To overcome the possible computational burden of implementing a large number of observers, we propose a dynamic sampling policy of the parameter set, which is inspired by the quantization strategy used for control in [27]. The idea is to update the sampling of the parameter set iteratively via a zoom-in procedure. With this dynamic scheme, we show that the parameter and state estimation errors also converge to the origin up to an adjustable margin with the added benefit of typically requiring less observers, as illustrated on an example.

The supervisory framework has traditionally been used for control, see [8], [16], [18], [30], [39]. In these works, the system dynamics depend on some unknown parameters and the objective is to steer the system state to the origin, however no guarantee is provided on the parameter estimates. A similar approach to the one we propose is pursued for the estimation of linear systems in [2], [3], [14], [4, Section 8.5], [26], where multiple observers are employed with different selection criteria. It has to be noted that the idea of dynamically sampling the parameter set has been used in [14] for linear systems under the assumption that the full state is measured, which is not the case in this study. In addition, we address nonlinear systems and we envision a different methodology compared to [14].

We believe that this paper illustrates the potential of casting the problem of estimation of

continuous-time systems in the hybrid systems setting. While hybrid tools have proved their efficiency for numerous control problems (see for example [12], [32], [34] to name a few), few works have investigated the estimation problem from this angle, see [1], [15] (for parameter estimation) and [22] (for state estimation). To the best of our knowledge, this paper is the first to address the joint estimation of states and parameters from this perspective. We believe that the advantages of supervisory control mentioned in [17] translate well to estimation. Firstly, it is not necessary to construct an adaptive observer, which is very challenging for nonlinear systems, see *e.g.* [9], [11], [42], [37]. Secondly, the supervisory framework has the advantage of modularity. Each of the components, namely the selection criterion, the monitoring signals and the multi-observer, can be designed independently to satisfy the respective properties required to meet our objective. This allows for the usage of readily available state observers for the additional purpose of parameter estimation. We thus show that Luenberger observers and a robust form of circle criterion observers [6], [10] can be used within this framework, noting that the approach is applicable to other types of state observers.

This work is motivated by the great need for estimation in neuroscience, especially for developing new methods for the classification or diagnosis of neurological diseases, such as epilepsy [36]. To illustrate the applicability of the framework, we implement the proposed algorithm to estimate the mean membrane potentials (states) and the synaptic gains (parameters) of a neural mass model [20] which is able to realistically reproduce patterns seen in electroencephalographic (EEG) recordings.

The paper is organised as follows. In Section II, we introduce the mathematical notation used in the paper. The problem is stated in Section III and we describe the supervisory setup in Section IV for the static sampling policy. The dynamic sampling policy is presented in Section V. The results obtained for both policies are shown to be applicable to linear systems and a class of nonlinear systems in Section VI. The illustrative example from neuroscience is presented in Section VII. Lastly, Section VIII concludes the paper. All proofs are provided in the Appendix.

II. PRELIMINARIES

Let $\mathbb{R} = (-\infty, \infty)$, $\mathbb{R}_{\geq 0} = [0, \infty)$, $\mathbb{R}_{> 0} = (0, \infty)$, $\mathbb{N} = \{0, 1, 2, \dots\}$ and $\mathbb{N}_{\geq 1} = \{1, 2, \dots\}$. The notation (u, v) stands for $[u^T \ v^T]^T$, where $u \in \mathbb{R}^m$ and $v \in \mathbb{R}^n$. For a vector $x \in \mathbb{R}^n$, $|x|$ denotes the Euclidean norm of x and $|x|_{\infty}$ denotes the infinity norm of x , i.e. $|x|_{\infty} :=$

$\max\{|x_1|, \dots, |x_n|\}$ where $x = (x_1, \dots, x_n)$. Let $\text{diag}(a_1, \dots, a_n)$ stand for the diagonal matrix with real elements a_1, \dots, a_n . The maximum (minimum) eigenvalue of a real, symmetric matrix P is denoted $\lambda_{\max}(P)$ ($\lambda_{\min}(P)$). The symmetric block component of a symmetric matrix is denoted by \star . The notation \mathbb{I} stands for the identity matrix. The hypercube centered at $\xi \in \mathbb{R}^n$ of edge length $2r > 0$ is denoted by $\mathcal{H}(\xi, r) := \{x \in \mathbb{R}^n : |x - \xi|_\infty \leq r\}$. For any $\Delta > 0$, let the set of piecewise continuous functions from $\mathbb{R}_{\geq 0}$ to $\mathcal{H}(0, \Delta)$ be \mathcal{M}_Δ . The left-limit operator is denoted by $(\cdot)^-$. A continuous function $\alpha : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ is a class \mathcal{K} function, if it is strictly increasing and $\alpha(0) = 0$; additionally, if $\alpha(r) \rightarrow \infty$ as $r \rightarrow \infty$, α is a class \mathcal{K}_∞ function. A continuous function $\beta : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ is a class \mathcal{KL} function, if $\beta(\cdot, s)$ is a class \mathcal{K} function for each $s \geq 0$ and $\beta(r, \cdot)$ is non-increasing and $\beta(r, s) \rightarrow 0$ as $s \rightarrow \infty$ for each $r \geq 0$.

III. PROBLEM FORMULATION

Consider the system

$$\begin{aligned}\dot{x} &= f(x, p^\star, u) \\ y &= h(x, p^\star),\end{aligned}\tag{1}$$

where the state is $x \in \mathbb{R}^{n_x}$, the measured output is $y \in \mathbb{R}^{n_y}$, the input is $u \in \mathbb{R}^{n_u}$ which is known and the *unknown* parameter vector $p^\star \in \Theta \subset \mathbb{R}^{n_p}$ is constant, where Θ is a known compact set. For any initial condition and any piecewise-continuous input u , system (1) admits a unique solution that is defined for all positive time. We make the following assumption on system (1).

Assumption 1. *The solutions to system (1) are uniformly bounded, i.e. for all $\Delta_x, \Delta_u \geq 0$, there exists a constant $K_x = K_x(\Delta_x, \Delta_u) > 0$ such that for all $x(0) \in \mathcal{H}(0, \Delta_x)$ and $u \in \mathcal{M}_{\Delta_u}$*

$$|x(t)|_\infty \leq K_x, \quad \forall t \geq 0.\tag{2}$$

□

It has to be noted that the bound K_x in (2) does not need to be known to implement the estimation algorithms presented in Sections IV and V, only its existence has to be ensured. Furthermore, Assumption 1 can be relaxed in some cases as explained later in Remark 2.

Remark 1. *Contrary to the problem of supervisory control in [39], we do not require system (1) to be stabilisable. Since our purpose is to estimate, we only require the solutions to system*

(1) to be uniformly bounded. This is a reasonable assumption in the estimation context since most physical systems have solutions that are uniformly bounded, as illustrated by the neural mass models investigated in Section VII.

The main objective of this paper is to estimate the parameter vector p^* and the state x of system (1) when only the input u and the output y of system (1) are measured, using a supervisory observer.

IV. SUPERVISORY OBSERVER WITH A STATIC SAMPLING POLICY

A. Description

Inspired by the supervisory framework used for control [28, Chapter 6], the proposed methodology consists of two basic units (see Figure 1): a bank of state observers (multi-observer) which generates state estimates and the supervisor (monitoring signals and a selection criterion) which chooses one observer at any given time. The estimated parameters and states are derived from the choice the supervisor makes.

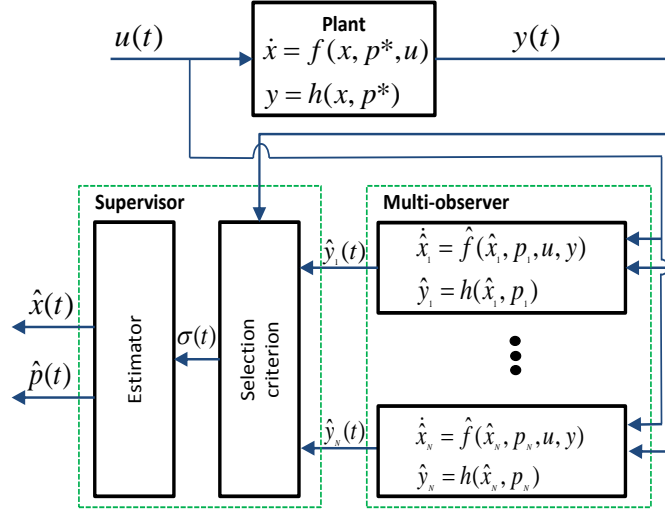


Fig. 1. Supervisory observer.

1) *Static sampling of the parameter set Θ* : We select $N \in \mathbb{N}_{\geq 1}$ parameter values p_i , $i \in \{1, \dots, N\}$, in the set Θ to form the sampled parameter set $\hat{\Theta} := \{p_1, \dots, p_N\}$. The selection

of the samples is done in such a way that the distance from p^* to $\hat{\Theta}$ defined as

$$d(p^*, \hat{\Theta}) := \min_{p \in \hat{\Theta}} |p^* - p|_\infty, \quad (3)$$

tends to zero as N tends to infinity. This can be guaranteed by embedding the parameter set Θ in a hyperrectangle (which is always possible because Θ is compact) and by employing uniform sampling, for instance. It may also be achieved using a logarithmic sampling if prior information is known, such as the probability distribution of the system parameter in Θ . Figure 2 illustrates these forms of sampling.

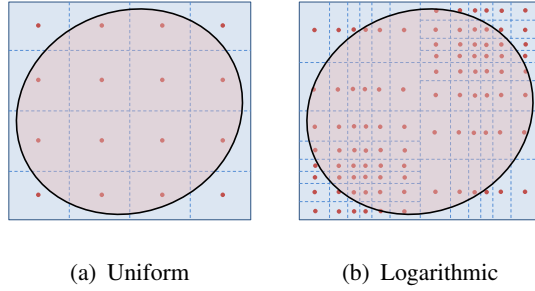


Fig. 2. Examples for the sampling of the parameter set $\Theta \subset \mathbb{R}^2$. Dots represent the parameters $p_i \in \hat{\Theta}$, $i \in \{1, \dots, N\}$. The shaded region in red indicates the parameter set Θ that has been embedded in a hyperrectangle (shaded region in blue).

2) *Multi-observer*: A state observer is designed for each p_i , $i \in \{1, \dots, N\}$,

$$\begin{aligned} \dot{\hat{x}}_i &= \hat{f}(\hat{x}_i, p_i, u, y) \\ \hat{y}_i &= h(\hat{x}_i, p_i), \end{aligned} \quad (4)$$

where $\hat{x}_i \in \mathbb{R}^{n_x}$ is the state estimate and $\hat{y}_i \in \mathbb{R}^{n_y}$ is the output estimate. The solutions to (4) are assumed to be unique and defined for all positive time for all initial conditions, any piecewise-continuous input u , any system output y and any parameter $p_i \in \Theta$. Denoting the state estimation error as $\tilde{x}_i := \hat{x}_i - x$, the output error as $\tilde{y}_i := \hat{y}_i - y$ and the parameter error as $\tilde{p}_i := p_i - p^*$, we obtain the following state estimation error systems, for $i \in \{1, \dots, N\}$,

$$\begin{aligned} \dot{\tilde{x}}_i &= \hat{f}(\tilde{x}_i + x, \tilde{p}_i + p^*, u, y) - f(x, p^*, u) =: F_i(\tilde{x}_i, x, \tilde{p}_i, p^*, u) \\ \tilde{y}_i &= h(\tilde{x}_i + x, \tilde{p}_i + p^*) - h(x, p^*) =: H(\tilde{x}_i, x, \tilde{p}_i, p^*). \end{aligned} \quad (5)$$

We assume that the observers (4) are designed such that the following property holds.

Assumption 2. *There exist scalars $a_1, a_2, \lambda_0 > 0$ and a continuous non-negative function $\tilde{\gamma} : \mathbb{R}^{n_p} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \rightarrow \mathbb{R}_{\geq 0}$ with $\tilde{\gamma}(0, x, u) = 0$ for all $x \in \mathbb{R}^{n_x}, u \in \mathbb{R}^{n_u}$ such that for any $p_i \in \Theta, i \in \{1, \dots, N\}$, there exists a continuously differentiable function $V_i : \mathbb{R}^{n_x} \rightarrow \mathbb{R}_{\geq 0}$, which satisfies the following for all $\tilde{x}_i \in \mathbb{R}^{n_x}, x \in \mathbb{R}^{n_x}, u \in \mathbb{R}^{n_u}$,*

$$a_1 |\tilde{x}_i|_\infty^2 \leq V_i(\tilde{x}_i) \leq a_2 |\tilde{x}_i|_\infty^2, \quad (6)$$

$$\frac{\partial V_i}{\partial \tilde{x}_i} F_i(\tilde{x}_i, \tilde{p}_i, p^*, u, x) \leq -\lambda_0 V_i(\tilde{x}_i) + \tilde{\gamma}(\tilde{p}_i, x, u). \quad (7)$$

□

Assumption 2 means that we know how to design a state observer for system (1) which is robust to parameter errors. When $\tilde{p}_i = 0$, i.e. when p^* is known, inequalities (6) and (7) imply that the origin of the state estimation error system (5) is globally exponentially stable (note that (6) can be equivalently stated with Euclidean norms as all norms are equivalent in \mathbb{R}^{n_x}). When $\tilde{p}_i \neq 0$, condition (7) needs to be satisfied, which is the case when system (5) is input-to-state exponentially stable [13] with respect to \tilde{p}_i for example. In Section VI, we will show that Luenberger observers and a class of nonlinear observers based on the circle criterion [6], [10] satisfy Assumption 2.

Following the terminology used in [39], we call the bank of observers in (4) the multi-observer.

Remark 2. *When Assumption 2 holds with $\tilde{\gamma}$ which only depends on \tilde{p}_i and u , we do not need system (1) to satisfy Assumption 2 to guarantee the convergence of the estimates.*

Remark 3. *The Lyapunov-based conditions (6)-(7) stated in Assumption 2 differ from the conditions in [39, Equations (10a) and (10b)] and [38, Theorem 4.3 (iii)] because the objective is to estimate the parameters and the states of (1), whereas the available results consider the problem of stabilization of an equilibrium point without guarantees on the convergence of the parameter estimates.*

3) *Monitoring signal:* Similar to [39, Equation 6], the monitoring signal associated with each observer is the exponentially weighted \mathcal{L}_∞ norm [19] of the output error defined as, for $i \in \{1, \dots, N\}$,

$$\mu_i(t) = \int_0^t \exp(-\lambda(t-s)) |\tilde{y}_i(s)|_\infty^2 ds, \quad \forall t \geq 0, \quad (8)$$

where $\lambda > 0$ is a design parameter. The monitoring signal (8) can be implemented as a linear filter, for $i \in \{1, \dots, N\}$,

$$\begin{aligned}\dot{\mu}_i(t) &= -\lambda\mu_i(t) + |\tilde{y}_i(t)|_\infty^2 & \forall t \geq 0 \\ \mu_i(0) &= 0.\end{aligned}\tag{9}$$

We assume that the output error of each of the observers \tilde{y}_i satisfies the following property.

Assumption 3. For any $\Delta_{\tilde{x}}, \Delta_x, \Delta_u > 0$, there exist a class \mathcal{K}_∞ function $\alpha_{\tilde{y}}$ and a constant $T_f = T_f(\Delta_{\tilde{x}}, \Delta_x, \Delta_u) > 0$ such that for all $p_i \in \Theta$, with $i \in \{1, \dots, N\}$, $\tilde{x}_i(0) \in \mathcal{H}(0, \Delta_{\tilde{x}})$, $x(0) \in \mathcal{H}(0, \Delta_x)$, for some $u \in \mathcal{M}_{\Delta_u}$, the corresponding solution to systems (1) and (5) satisfies

$$\int_{t-T_f}^t |\tilde{y}_i(\tau)|_\infty^2 d\tau \geq \alpha_{\tilde{y}}(|\tilde{p}_i|_\infty), \quad \forall t \geq T_f.\tag{10}$$

□

The inequality (10) is known as a *persistence of excitation (PE) condition* that appears in the identification and adaptive literature [35]. It differs from the classical PE definition [35, Definition 2.5.3] in that we take the norm of the signal and we consider a family of systems (5) parameterised by \tilde{p}_i , where we consider only the lower bound (excitation level) in (10) which depends on \tilde{p}_i . In particular, if $\tilde{p}_i = 0$, we do not require any PE property. In addition, the excitation level grows with the norm of the parameter error \tilde{p}_i . Hence, the integral term in (10) provides quantitative information about the parameter estimation error. Assumption 3 holds when the output errors \tilde{y}_i , $i \in \{1, \dots, N\}$, satisfy the classical PE condition according to the proposition below.

Proposition 1. Consider system (1) and the observer (4). Suppose that the following holds.

- 1) Assumption 1 is satisfied.
- 2) The functions f and h in (1) and \hat{f} in (4) are continuously differentiable.
- 3) For all $\Delta_{\tilde{x}}, \Delta_x, \Delta_u > 0$ and any $p_i \in \Theta$, $i \in \{1, \dots, N\}$, with $\tilde{p}_i \neq 0$ there exist constants $T_f = T_f(\Delta_{\tilde{x}}, \Delta_x, \Delta_u)$ and $\bar{\alpha}_i = \bar{\alpha}_i(\Delta_{\tilde{x}}, \Delta_x, \Delta_u, \tilde{p}_i) > 0$ such that for all $\tilde{x}_i(0) \in \mathcal{H}(0, \Delta_{\tilde{x}})$, $x(0) \in \mathcal{H}(0, \Delta_x)$ and for some $u \in \mathcal{M}_{\Delta_u}$, the corresponding solution to (1), (5) satisfies

$$\int_{t-T_f}^t \tilde{y}_i(\tau) \tilde{y}_i(\tau)^T d\tau \geq \bar{\alpha}_i \mathbb{I}, \quad \forall t \geq T_f.\tag{11}$$

Then (10) holds.

□

Proposition 1 indicates that tools to verify the classical PE condition (11) can be applied to ensure the satisfaction of Assumption 3. The reader is referred to [31, Chapter 6] and [33] where a priori checkable conditions are respectively given for linear and nonlinear systems.

4) *Selection criterion:* The signal $\sigma : \mathbb{R}_{\geq 0} \rightarrow \{1, \dots, N\}$ is used to choose an observer from the bank of N observers at every instant of time. It is defined as

$$\sigma(t) := \arg \min_{i \in \{1, \dots, N\}} \mu_i(t), \quad \forall t \geq 0. \quad (12)$$

Note that no dwell time is guaranteed with this selection criterion, therefore rapid changes of the signal σ are allowed¹. Nevertheless, these switches do not affect the dynamics of the observers (4) and system (1).

5) *Parameter and state estimates:* Based on the signal σ in (12), the estimated parameters and states respectively are, for all $t \geq 0$,

$$\hat{p}(t) := p_{\sigma(t)}, \quad (13)$$

$$\hat{x}(t) := \hat{x}_{\sigma(t)}(t). \quad (14)$$

The parameter and the state estimates are discontinuous in general because these signals switch among a finite family of continuous trajectories that are in general different at the switching instant.

B. Convergence guarantees

The theorem below states that the estimated parameter \hat{p} and state \hat{x} in (13) and (14) are respectively guaranteed to converge to their true values p^* and x up to some selected margins $\nu_{\hat{p}}$ and $\nu_{\hat{x}} > 0$, provided that the number of observers N is sufficiently large.

Theorem 1. *Consider system (1), the multi-observer (4), the monitoring signals (8), the selection criterion (12), the parameter estimate (13) and the state estimate (14). Suppose Assumptions 1-3 are satisfied. For any $\Delta_{\tilde{x}}, \Delta_x, \Delta_u > 0$ and any margins $\nu_{\tilde{x}}, \nu_{\hat{p}} > 0$, there exist $T, \bar{K}_{\tilde{x}} > 0$*

¹The hysteresis based switching in [18] can be used instead of (12) to ensure the existence of a finite number of discontinuities of σ over any given finite-time interval. The results of Sections IV-B and V-B still apply in this case provided that a positive constant is added to the monitoring signal as done in [18, Equation 9] and that the hysteresis constant is sufficiently small.

and a sufficiently large $N^* \in \mathbb{N}_{\geq 1}$ such that for any $N \geq N^*$, the following holds for all $(x(0), \tilde{x}_i(0)) \in \mathcal{H}(0, \Delta_x) \times \mathcal{H}(0, \Delta_{\tilde{x}})$ for $i \in \{1, \dots, N\}$ and for any $u \in \mathcal{M}_{\Delta_u}$ that satisfies Assumption 3,

$$|\tilde{p}_{\sigma(t)}(t)|_{\infty} \leq \nu_{\tilde{p}} \quad \forall t \geq T, \quad \limsup_{t \rightarrow \infty} |\tilde{x}_{\sigma(t)}(t)|_{\infty} \leq \nu_{\tilde{x}}, \quad |\tilde{x}_{\sigma(t)}(t)|_{\infty} \leq \bar{K}_{\tilde{x}} \quad \forall t \geq 0. \quad (15)$$

□

The accuracy of the estimates can be rendered as accurate as desired by increasing the number of observers N . To ensure the properties in (15), what we actually need is that a selected parameter p_i is sufficiently close to p^* so that $d(p^*, \hat{\Theta})$ in (3) is sufficiently small. This is guaranteed by taking N to be large with a sampling of Θ which ensures (3). For some applications, it is sufficient to have estimates that are accurate within some margin of error. This is the case for instance for the anticipation of abnormal neural behaviour such as seizures caused by epilepsy, where the model parameter set is composed of seizure and non-seizure related subsets [40]. Therefore, an algorithm that provides estimates of the parameters within some adjustable margin can be used to infer the region of the parameter set in which the true parameter lies.

A potential drawback of the scheme presented in this section is the need for a sufficiently large number of observers to ensure that the estimates fall within the required margins. While this may be feasible for some applications, it may be computationally intensive for others. We explain how to overcome this issue in the next section.

V. SUPERVISORY OBSERVER WITH A DYNAMIC SAMPLING POLICY

Contrary to Section IV, we now update the sampling of the parameter set Θ at some time t_k , $k \in \mathbb{N}$. The proposed dynamic sampling policy builds upon the results of Section IV in that with a sufficient number of observers N , the parameter estimation error converges to a given margin in finite-time according to Theorem 1. Once this happens, a new set of N parameters is chosen within the hypercube centered at the latest parameter estimate with an edge length which is proportional to the aforementioned margin. This aims to provide a better estimate of the parameter and the state for a given number of observers by zooming into the parameter set at each update time.

A. Description

For the sake of simplicity and without loss in generality, we assume that Θ is a hypercube² centered at some known $p_c \in \mathbb{R}^{n_p}$ and of edge length $2\Delta > 0$. Let $N \in \mathbb{N}_{\geq 1}$ be the number of parameter samples and $\nu_{\tilde{p}} > 0$ be the acceptable margin of error for the parameter estimate.

At the initial time $t_0 = 0$, we select N values in Θ to form the initial sampled parameter set $\hat{\Theta}(0) := \{p_1(0), \dots, p_N(0)\}$. The way the set $\hat{\Theta}(0)$ is defined is as explained in the following. From the results of the static policy in Section IV (in particular (15)), we know that after a sufficiently long time T (see (15)) and for a sufficiently large N , the norm of the parameter estimation error will be less than $C := \max\{\alpha\Delta, \nu_{\tilde{p}}\}$ where $\alpha \in (0, 1)$ is a design constant, which we call the zooming factor. In the case where $C = \alpha\Delta$, we know that p^* is in the hypercube $\mathcal{H}(\hat{p}(t_1^-), \alpha\Delta)$ for $t_1 \geq T$. We can thus select N parameter values in the hyperrectangle $\Theta(1) := \mathcal{H}(\hat{p}(t_1^-), \alpha\Delta) \cap \Theta$ (noting that $\mathcal{H}(\hat{p}(t_1^-), \alpha\Delta)$ is not necessarily included in Θ) to form a new set of sampled parameters $\hat{\Theta}(1) := \{p_1(1), \dots, p_N(1)\}$, see Figure 3. When $C = \nu_{\tilde{p}}$, the desired result is obtained. Nevertheless, we cannot detect this case on-line as we do not know \tilde{p}_σ . Hence, we apply the same procedure as when $C = \alpha\Delta$, even though p^* is no longer guaranteed to be in $\mathcal{H}(\hat{p}(t_1^-), \alpha\Delta)$, which is fine as the norm of the induced parameter estimation error will remain less than $\nu_{\tilde{p}}$ for all future times. This procedure is carried out iteratively at each update time instant t_k , $k \in \mathbb{N}$ which verifies

$$T_d := t_{k+1} - t_k > 0, \quad \forall k \in \mathbb{N}, \quad (16)$$

where $T_d > 0$ is a design parameter which is selected larger than T . The sets $\hat{\Theta}(k)$, for $k \in \mathbb{N}$, are defined such that the property below is verified

$$p^* \in \Theta(k) \Rightarrow d(p^*, \hat{\Theta}(k)) \leq \pi(\Delta, N), \quad (17)$$

where $\pi \in \mathcal{KL}$. In that way, for any $k \in \mathbb{N}$, $d(p^*, \hat{\Theta}(k)) \rightarrow 0$ as $N \rightarrow \infty$ like in Section IV (as long as $p^* \in \Theta(k)$). An example of sampling which ensures (17) is provided below.

²We can always find a hypercube which contains Θ (since Θ is a compact set) and take the hypercube to be the parameter set. It has to be noted that to embed a given parameter set into a hypercube may be conservative, as we may end up working with a ‘much larger set’ compared to the initial one. In this case, the parameters may be scaled to reduce this conservatism.

Example 1. It can be noted that the sets $\Theta(k)$, $k \in \mathbb{N}$, constructed in this section are hyperrectangles, i.e.

$$\Theta(k) = \{p = (p^1, \dots, p^{n_p}) : \forall j \in \{1, \dots, n_p\} |p^j - p_c^j(k)| \leq \delta^j(k)\}, \quad (18)$$

where $p_c(k) = (p_c^1(k), \dots, p_c^{n_p}(k))$ is the center and the constants $\delta^j(k) > 0$, $j \in \{1, \dots, n_p\}$, define the edge lengths. Moreover it holds that $\Theta(k) \subset \Theta$, hence $\delta^j(k) \leq \Delta$ for any $j \in \{1, \dots, n_p\}$. Let $N = m^{n_p}$ with $m \in \mathbb{N}_{\geq 1}$, we partition $\Theta(k)$ into N hyperrectangles such that their union is equal to $\Theta(k)$ and the intersection of their interiors is empty. These N sets have the form $\{p = (p^1, \dots, p^{n_p}) : \forall j \in \{1, \dots, n_p\} |p^j - p_{c,i}^j(k)| \leq \frac{\delta^j(k)}{m}\}$ where $p_{c,i}$ is the center, $i \in \{1, \dots, N\}$. Suppose $p^* \in \Theta(k)$, then $p^* = (p_1^*, \dots, p_{n_p}^*)$ belongs to (at least) one of these N hyperrectangles. As a consequence, $|p_j^* - p_{c,i^*}^j(k)| \leq \frac{\delta^j(k)}{m}$ for some $i^* \in \{1, \dots, N\}$ and any $j \in \{1, \dots, n_p\}$. Consequently $|p^* - p_{c,i^*}|_\infty \leq \max_{j \in \{1, \dots, n_p\}} \frac{\delta^j(k)}{m} \leq \frac{\Delta}{m}$ as $\delta^j(k) \leq \Delta$ for any $j \in \{1, \dots, n_p\}$. Finally, we derive that $d(p^*, \hat{\Theta}(k)) \leq |p^* - p_{c,i^*}|_\infty \leq \frac{\Delta}{N^{\frac{1}{n_p}}}$. Noting that $\pi : (s, r) \mapsto \min \left\{ s, \frac{s}{r^{\frac{1}{n_p}}} \right\}$, we have that (17) holds. \square

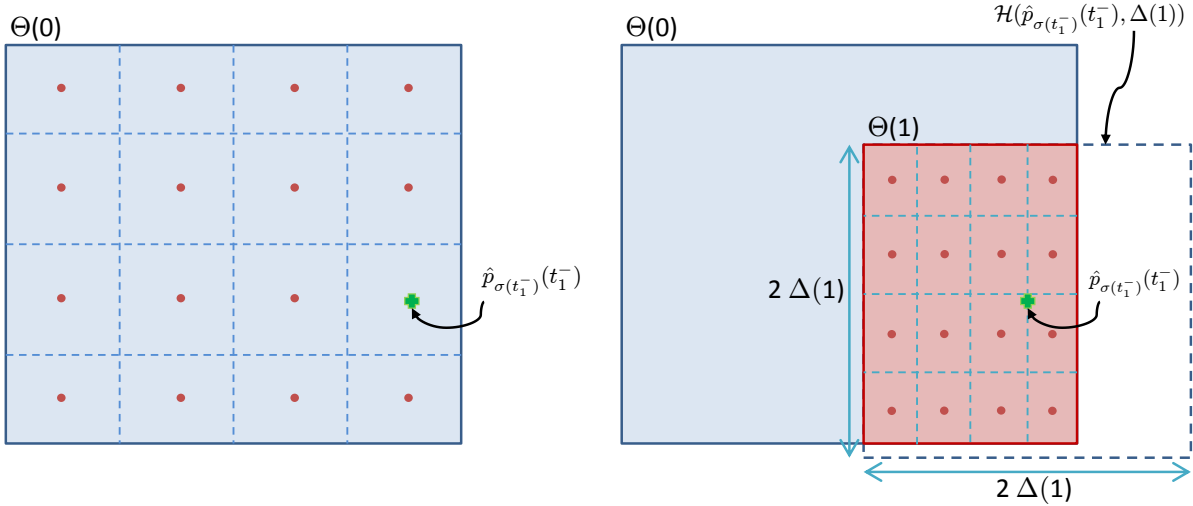


Fig. 3. Illustration of sets $\Theta(0)$ (shaded blue region), $\Theta(1)$ (shaded red region), $\hat{\Theta}(0)$ and $\hat{\Theta}(1)$. Legend: red dots indicate the sampled parameters p_i for $i \in \{1, \dots, N\}$, green cross indicates the selected parameter $\hat{p}_{\sigma(t_1^-)}(t_1^-)$.

We summarize the dynamic sampling policy below.

- At $t_0 = 0$. Let $\Delta(0) = \Delta$ and $\Theta(0) = \Theta$. The set $\hat{\Theta}(0)$ is obtained by discretizing the set $\Theta(0)$ with N points such that (17) holds.

- At $t = t_k$ for $k \in \mathbb{N}$. Let

$$\Delta(k) = \alpha \Delta(k-1), \quad (19)$$

where the zooming factor $\alpha \in (0, 1)$ is a design parameter. We define the zoomed-in parameter set $\Theta(k)$ as

$$\Theta(k) := \mathcal{H}(\hat{p}(t_k^-), \Delta(k)) \cap \Theta(k-1) \cap \dots \cap \Theta(0). \quad (20)$$

The sampled parameter set $\hat{\Theta}(k) := \{p_1(k), \dots, p_N(k)\}$ consists of N points which are selected such that (17) is verified.

The dynamic sampling policy leads to the following changes in two components of the supervisory observer, namely the multi-observer (4) and the implementation of the monitoring signals (8) described in Section IV. In the multi-observer, a state observer is designed for each sampled parameter $p_i(t)$, $i \in \{1, \dots, N\}$ for $t \in [t_k, t_{k+1})$,

$$\begin{aligned} \dot{\hat{x}}_i &= \hat{f}(\hat{x}_i, p_i, u, y), & \forall t \in [t_k, t_{k+1}), k \in \mathbb{N}, \\ \hat{y}_i &= h(\hat{x}_i, p_i), \\ \hat{x}_i(t_k) &= \hat{x}_i(t_k^-), \end{aligned} \quad (21)$$

where $\hat{x}_i \in \mathbb{R}^{n_x}$ is the state estimate provided by observer i . Note that the state \hat{x}_i does not jump at update times t_k , $k \in \mathbb{N}$. Secondly, the definition of the monitoring signals (8) remains unchanged, but the implementation as linear filters become, for $i \in \{1, \dots, N\}$,

$$\begin{aligned} \dot{\mu}_i(t) &= -\lambda \mu_i(t) + |\tilde{y}_i(t)|_\infty^2, & \forall t \in [t_k, t_{k+1}), k \in \mathbb{N}, \\ \mu_i(t_k) &= 0. \end{aligned} \quad (22)$$

B. Convergence guarantees

We are now ready to state the main result of this section.

Theorem 2. *Consider system (1), the multi-observer (21), the monitoring signals (22), the selection criterion (12), the parameter estimate (13), the state estimate (14) and the dynamic sampling policy (16)-(20). Suppose Assumptions 1-3 are satisfied. For any $\Delta_{\tilde{x}}, \Delta_x, \Delta_u > 0$, any margins $\nu_{\tilde{x}}, \nu_{\tilde{p}} > 0$ and zooming factor $\alpha \in (0, 1)$, there exist $\bar{T}, \bar{K}_{\tilde{x}} > 0$, sufficiently large $T^* > 0$ and $N^* \in \mathbb{N}_{\geq 1}$ such that for any $N \geq N^*$ and $T_d \geq T^*$, for all $(x(0), \tilde{x}_i(0)) \in \mathcal{H}(0, \Delta_x) \times \mathcal{H}(0, \Delta_{\tilde{x}})$ for $i \in \{1, \dots, N\}$ and for any $u \in \mathcal{M}_{\Delta_u}$ that satisfies Assumption 3, the following holds*

$$|\tilde{p}_{\sigma(t)}(t)|_\infty \leq \nu_{\tilde{p}} \quad \forall t \geq \bar{T}, \quad \limsup_{t \rightarrow \infty} |\tilde{x}_{\sigma(t)}(t)|_\infty \leq \nu_{\tilde{x}}, \quad |\tilde{x}_{\sigma(t)}(t)|_\infty \leq \bar{K}_{\tilde{x}}, \quad \forall t \geq 0. \quad (23)$$

□

Theorem 2 states that the estimated parameters \hat{p} and states \hat{x} in (13) and (14), are respectively ensured to converge to their true values p^* and x within some selected margins $\nu_{\hat{p}}$ and $\nu_{\hat{x}}$. These guarantees are the same as those in Theorem 1. Nevertheless, for a given set of N observers, Theorem 2 may only ensure a better accuracy of the estimates compared to Theorem 1. Indeed, consider a number of observers N (sufficiently large). On the interval $[0, t_1]$, the static and the dynamic schemes provide the same estimates. Then, at time t_1 , the dynamic scheme will start the zoom-in procedure which may only reduce the estimation error on the parameter and thus on the state. Hence, the dynamic scheme may be used to reduce the number of observers needed to ensure the convergence of the parameter and state estimation error up to given margins, as illustrated in Section VII.

VI. APPLICATIONS

In this section, we apply the results of Theorem 2 to two case studies: stable linear systems and a class of nonlinear systems.

A. Linear systems

We consider the linear system

$$\begin{aligned}\dot{x} &= A(p^*)x + B(p^*)u \\ y &= C(p^*)x,\end{aligned}\tag{24}$$

where $x \in \mathbb{R}^{n_x}$, $u \in \mathbb{R}^{n_u}$, $y \in \mathbb{R}^{n_y}$, $p^* \in \Theta \subset \mathbb{R}^{n_p}$. The matrices $A(p)$, $B(p)$ and $C(p)$ are continuous in p on Θ , the pair $(A(p), C(p))$ is detectable for any $p \in \Theta$ and $A(p^*)$ is Hurwitz, which ensures the satisfaction of Assumption 1.

Each observer in (4) is designed as follows for $p_i \in \Theta$, $i \in \{1, \dots, N\}$,

$$\begin{aligned}\dot{\hat{x}}_i &= A(p_i)\hat{x}_i + B(p_i)u + L(p_i)(C(p_i)\hat{x}_i - y) \\ \hat{y}_i &= C(p_i)\hat{x}_i,\end{aligned}\tag{25}$$

where $L(p_i)$ is such that $A(p_i) + L(p_i)C(p_i)$ is Hurwitz (this is always possible since $(A(p_i), C(p_i))$ is detectable). The proposition below shows that Assumption 2 is satisfied.

Proposition 2. *Consider the linear system (24) and the observer (25). Assumption 2 holds. \square*

If the classical PE condition (11) is guaranteed, Assumption 3 is satisfied according to Proposition 1 (note that item 2 of Proposition 1 holds for the considered system). There exist results in the literature (see Chapter 6 in [31]) which provide sufficient conditions to verify (11) for linear systems as mentioned earlier. These results can be used, for instance, to design an input u to system (24) such that the inequality (11) is satisfied. The proposition below directly follows from Proposition 1 and Theorems 1-2.

Proposition 3. *Consider system (24) and state-observer (25) and suppose (11) holds. When the static sampling policy described in Section IV (respectively the dynamic sampling policy of Section V) is used, then the conclusions of Theorem 1 (respectively Theorem 2) hold. \square*

B. A class of nonlinear systems

We consider the following class of nonlinear systems studied in [7]

$$\begin{aligned}\dot{x} &= A(p^*)x + G(p^*)\gamma(Hx) + B(p^*)\phi(u, y) \\ y &= C(p^*)x,\end{aligned}\tag{26}$$

where $x \in \mathbb{R}^{n_x}$, $u \in \mathbb{R}^{n_u}$, $y \in \mathbb{R}^{n_y}$, $p^* \in \Theta \subset \mathbb{R}^{n_p}$, $\gamma : \mathbb{R}^{n_p} \rightarrow \mathbb{R}^{n_\gamma}$ and $\phi : \mathbb{R}^{n_u} \times \mathbb{R}^{n_y} \rightarrow \mathbb{R}^{n_\sigma}$. The matrices $A(p)$, $G(p)$ and $B(p)$ are continuous in p on Θ . We assume that Assumption 1 holds which can be verified with the aid of Lyapunov-based tools [24]. We now explain how to design a state-observer which ensures the satisfaction of Assumption 2. For that purpose, we make the following assumption on the vector field γ , like in [7], [10], [41].

Assumption 4. *For any $k \in \{1, \dots, n_\gamma\}$, there exist constants $a_{\gamma_k} \in \mathbb{R}$, $b_{\gamma_k} \in \mathbb{R} \setminus \{0\}$ such that the following holds*

$$-\infty < a_{\gamma_k} \leq \frac{\partial \gamma_k(v_k)}{\partial v_k} \leq b_{\gamma_k} < \infty, \quad \forall v_k \in \mathbb{R},\tag{27}$$

where $\gamma = (\gamma_1, \dots, \gamma_{n_\gamma})$. \square

The following state-observer [7], [10], [41] is designed for any $p_i \in \Theta$, $i \in \{1, \dots, N\}$,

$$\begin{aligned}\dot{\hat{x}}_i &= A(p_i)\hat{x}_i + G(p_i)\gamma(H\hat{x}_i + K(p_i)(C\hat{x}_i - y)) + B(p_i)\phi(u, y) + L(p_i)(C(p_i)\hat{x}_i - y) \\ \hat{y}_i &= C(p_i)\hat{x}_i,\end{aligned}\tag{28}$$

where $K(p_i)$ and $L(p_i)$ are the observer matrices. These matrices are selected such that the inequality (29) below holds.

Proposition 4. *Consider system (26) and state-observer (28) for $i \in \{1, \dots, N\}$. Suppose the following holds.*

- 1) *Assumption 4 holds.*
- 2) *There exist real matrices $P_i = P_i^T > 0$, $M_i = \text{diag}(m_{i1}, \dots, m_{in_p}) > 0$ and scalars ν_i , $\mu_i > 0$ such that the following holds*

$$\begin{bmatrix} \mathcal{A}(P_i, L(p_i), \nu_i) & \mathcal{B}(P_i, M_i, K(p_i)) & P_i \\ \star & \mathcal{E}(M_i) & 0 \\ \star & \star & -\mu_i \mathbb{I} \end{bmatrix} \leq 0, \quad (29)$$

where $\mathcal{A}(P_i, L(p_i), \nu_i) = P_i(A(p_i) + L(p_i)C(p_i)) + (A(p_i) + L(p_i)C(p_i))^T P_i + \nu_i \mathbb{I}$,

$\mathcal{B}(P_i, M_i, K(p_i)) = P_i G(p_i) + (H + K(p_i)C(p_i))^T M_i$ and $\mathcal{E}(M_i) = -2M_i \text{diag}\left(\frac{1}{b_{\gamma_1}}, \dots, \frac{1}{b_{\gamma_{n_p}}}\right)$.

Then Assumption 2 is satisfied. \square

Note that inequality (29) is considered a linear matrix inequality (LMI) in P_i , $P_i L(p_i)$, $M_i K(p_i)$, M_i and μ_i . Therefore, (29) can be solved using efficient software such as the LMI solvers in MATLAB.

Remark 4. *Other than the circle criterion based observer (28), any exponential observer³ for which its convergence is ensured using a quadratic Lyapunov function also satisfy Assumption 3 provided that the right hand sides of (4) are continuous in the parameter p .*

We assume that the PE condition stated in Assumption 3 is satisfied. PE properties for nonlinear systems are studied in [33] and may be used in conjunction with Proposition 1 to verify Assumption 3. The following result is a direct application of Proposition 1 and Theorems 1-2.

Proposition 5. *Consider system (26), state-observer (28) and suppose the following holds.*

- 1) *Assumptions 1 and 4 and (11) hold.*

³An exponential observer has a corresponding state estimation error system whose equilibrium is exponentially stable when there is no parameter mismatch, i.e. $\tilde{p}_i = 0$.

2) Condition (29) is feasible.

When the static sampling policy described in Section IV (respectively the dynamic sampling policy of Section V) is employed, then the conclusions of Theorem 1 (respectively of Theorem 2) hold. \square

VII. ILLUSTRATIVE EXAMPLE: A NEURAL MASS MODEL

In this section, we apply the results of Section VI-B to estimate the synaptic gains (parameters) and the mean membrane potentials (states) of neuronal populations of the neural mass model in [20]. The model in [20] describes the dynamics of a single cortical column by capturing the interactions between the pyramidal neurons, the excitatory and the inhibitory interneurons in a localised region of the cortex. It has been shown to realistically reproduce various patterns seen in the EEG recordings which is the measured output, such as alpha rhythms. Moreover, it may be used to generate more complex phenomena as shown in [40].

To write the model in the form of (26), we take the states⁴ to be $x = (x_{01}, x_{02}, x_{11}, x_{12}, x_{21}, x_{22}) \in \mathbb{R}^6$, where x_{01}, x_{11} and x_{21} are the the membrane potential contributions of the pyramidal neurons, the excitatory and inhibitory interneurons respectively and x_{02}, x_{12} and x_{22} are their respective time-derivatives. The unknown vector of parameters⁴ is $p^* = (p_1^*, p_2^*)$, where p_1^* and p_2^* represent the synaptic gains of the excitatory and inhibitory neuronal populations respectively. The vector of parameters p^* belongs to $\Theta := [4, 8] \times [22, 28]$ in agreement with [20]. The matrices in (26) are defined as $C = \begin{pmatrix} 0 & 0 & 1 & 0 & -1 & 0 \end{pmatrix}$, $A = \text{diag}(A_a, A_a, A_b)$, where $A_a = \begin{pmatrix} 0 & 1 \\ -a^2 & -2a \end{pmatrix}$,

$$A_b = \begin{pmatrix} 0 & 1 \\ -b^2 & -2b \end{pmatrix}, G(p^*) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ p_1 a c_2 & 0 \\ 0 & 0 \\ 0 & p_2 b c_4 \end{pmatrix}, B(p^*) = \begin{pmatrix} 0 & 0 \\ p_1 a & 0 \\ 0 & 0 \\ 0 & p_1 a \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

and $H = \begin{pmatrix} c_1 & 0 & 0 & 0 & 0 & 0 \\ c_3 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$. The parameters a, b, c_1, c_2, c_3 and $c_4 \in \mathbb{R}_{>0}$ are assumed

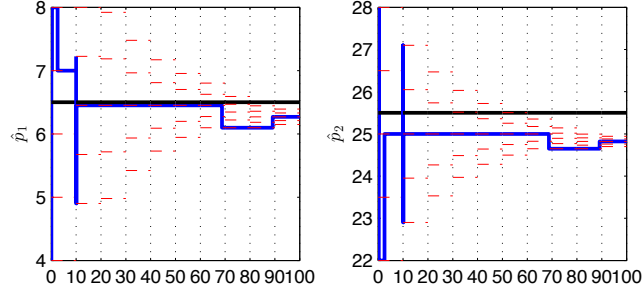
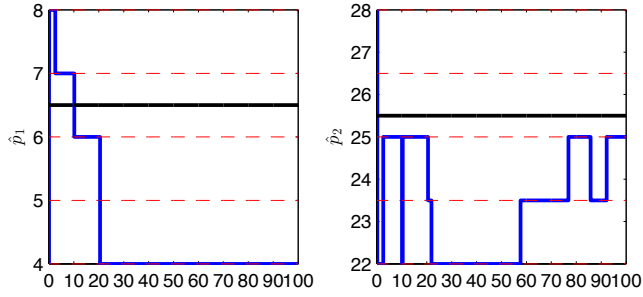
⁴According to the notation of [20], $x = (y_0, y_3, y_1, y_4, y_2, y_5)$ and the parameter vector is taken to be $p^* = (A, B)$.

to be known. The nonlinear terms in (26) are $\gamma = (S, S)$ and $\phi(u, y) = (S(y), u)$. The function S denotes the sigmoid function $S(v) := \frac{2e_0}{1+e^{r(v_0-v)}}$ for $v \in \mathbb{R}$, with known constants $e_0, v_0, r \in \mathbb{R}_{>0}$. For a detailed description of the model and its parameters, see [20]. The model has uniformly bounded solutions for all initial conditions and bounded input $u \in \mathcal{M}_{\Delta_u}$, because the matrix A is Hurwitz and the nonlinearity S in γ and ϕ is bounded. Therefore, the model satisfies Assumption 1. Furthermore, by the definition of the nonlinearity S above, Assumption 4 is satisfied.

We perform simulations with the model initialised at $x(0) = 10 \times 1_{6 \times 1}$, $p^* = (p_1^*, p_2^*) = (6.5, 25.5)$. Since the model can be written in the form of (26), we design the state-observers (4) of the form (28). The gains $K(p_i)$ and $L(p_i)$ are obtained by solving (29). We implement the supervisory observer with both the static and dynamic sampling policies of Sections IV and V. The following design parameters are chosen: $\lambda = 0.005$ (for the monitoring signal in Section IV-A3), the zooming factor $\alpha = 0.8$ and the sampling interval $T_d = 10$ s (for the dynamic sampling policy in Section V).

In Figures 4-5, we compare the performance of the supervisory observer with and without dynamic sampling of parameters for $N = N_A \times N_B = 5 \times 5$ observers, where N_A and N_B are the number of samples taken in the set of possible p_1^* and p_2^* values, respectively. Both policies allow us to estimate the states and the parameter up to some margin of the true values. The dynamic policy also gives better results in view of Figure 5. To illustrate the workings of the dynamic sampling policy, we provide snapshots of the sampled parameters for $t \in [0, 150]$ s at each update time t_k for $k \in \{0, 1, 5, 9, 10, 14\}$ in Figure 6. Note that for $t \geq 60$ s, the plant parameter is no longer in the zoomed parameter set $\Theta(6)$, a phenomenon that is expected (see Section V). Nevertheless, the parameter estimate has converged to a desired neighbourhood of the plant parameter p^* for $t \geq 60$ s.

To investigate the impact of the number of state-observers on the convergence results, we consider three values for $N = N_A \times N_B$: 2×2 , 4×4 and 5×5 . Table I summarises the results obtained. We see that the supervisory observer with the dynamic policy outperforms the static scheme, in terms of the accuracy of the parameter estimate. This does not necessarily translate to a smaller ultimate bound on the state estimation error as seen in the results obtained for the static sampling policy. This can be explained as follows. Firstly, the results of Sections IV and V hold for a sufficiently large number of observers N and it may be the case that $N = 4$ and

(a) Dynamic sampling with sampling interval $t_{k+1} - t_k = 10s$.

(b) Static sampling

Fig. 4. Estimated parameter \hat{p} and the sampled parameter values. Legend: red dashed lines denote the chosen parameter values p_i , $i \in \{1, \dots, N\}$ and the vertical, dotted lines indicate the time of updating the parameter samples, t_k , $k \in \mathbb{N}$. The black solid line indicates the system parameter values p^* .

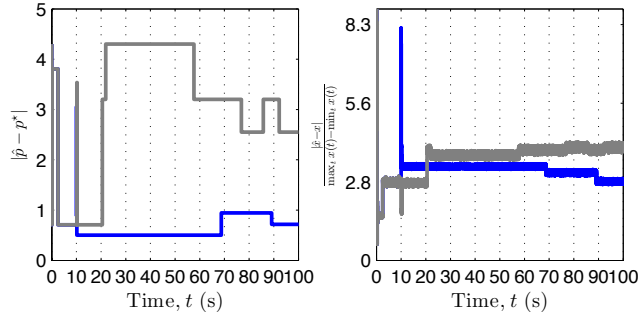


Fig. 5. Norms of the parameter and state estimation errors obtained using the dynamic sampling policy with sampling interval $T_d = 10s$ (blue line) and the static sampling policy (grey line).

$N = 16$ are not large enough to satisfy the conditions of Theorems 1 and 2. Secondly, although each observer satisfies Assumption 2, these state-observers do not have the same quantitative robustness properties with respect to the parameter estimation error \tilde{p}_i . To be precise, $\bar{\gamma}_{\tilde{x}}$ of the individual state estimation error systems (5) are in general different. Thus, the decrease of $|\tilde{p}_i|$

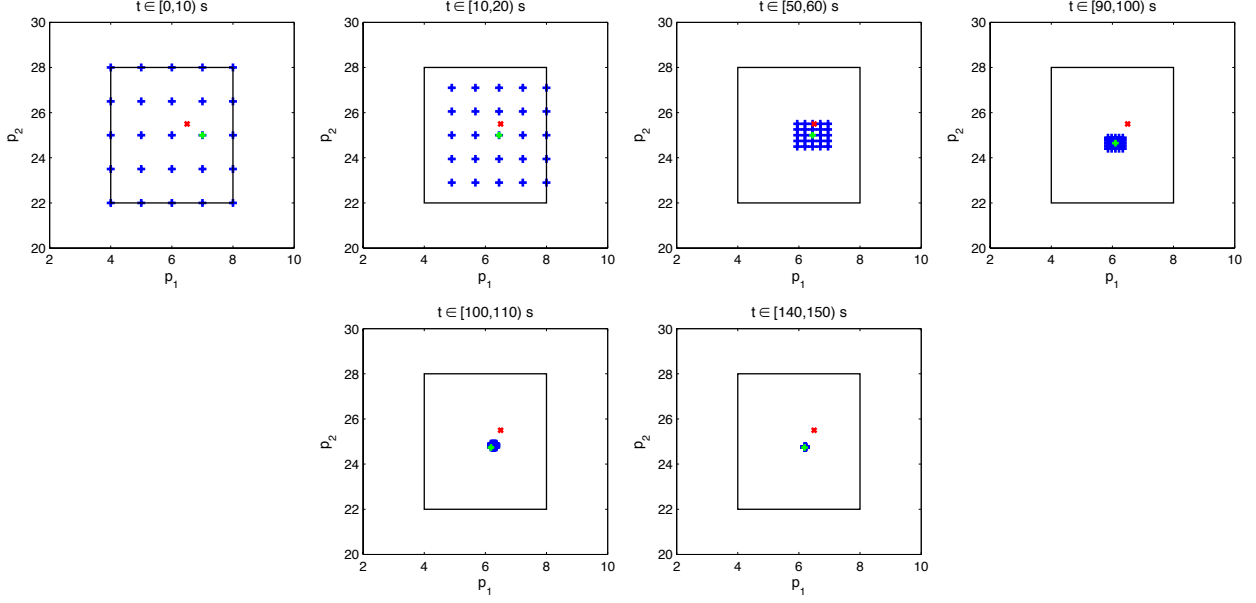


Fig. 6. The dynamic sampling of the parameter set at some update instant t_k for $(N_A, N_B) = (5, 5)$ with sampling interval $t_{k+1} - t_k = 10$ s. Legend: \times is the true parameter p^* ; $+$ are the sampled parameters; $+$ is the final selected parameter in the interval $t \in [t_k, t_{k+1})$; the known parameter set Θ is drawn with solid black lines.

may be compensated by larger $\bar{\gamma}_{\tilde{x}}$. It will therefore be interesting to develop observers which minimise $\bar{\gamma}_{\tilde{x}}$ in future work.

$N = N_A \times N_B$	2×2	4×4	5×5
Static policy: $ \tilde{p}_{\sigma(t_f)}(t_f) $	4.30	3.80	2.55
Dynamic policy: $ \tilde{p}_{\sigma(t_f)}(t_f) $	2.66	1.04	0.72
Static policy: $\frac{ \tilde{x}_{\sigma(t_f)}(t_f) }{\max_t x(t) - \min_t x(t) }$	4.01	1.73	4.64
Dynamic policy: $\frac{ \tilde{x}_{\sigma(t_f)}(t_f) }{\max_{t \in [0, t_f]} x(t) - \min_{t \in [0, t_f]} x(t) }$	4.26	3.47	3.22

The final time of simulation t_f is 100s.

TABLE I

NUMERICAL RESULTS FOR INCREASING NUMBER OF OBSERVERS N SUCH THAT $d(p^*, \hat{\Theta})$ DECREASES IN THE STATIC CASE.

VIII. CONCLUSIONS

The main contribution of this paper is the design of schemes for the parameter and state estimation of nonlinear continuous-time systems with convergence guarantees. The proposed

approach builds upon recent advances on supervisory control. It allows us to treat nonlinear systems provided robust state observers with respect to parameter errors can be designed and a given PE condition holds. We have shown that the parameter estimates converge in finite-time to the true parameter with any desired accuracy and that the norm of the state estimation error converges to the origin up to an adjustable margin, by taking a sufficiently large number of state-observers. This scheme may be computationally intense in some cases. This motivated the introduction of a dynamic sampling policy of the parameter set which may be used to ensure the same convergence guarantees with less number of observers. We have shown how these results can be applied to linear systems using Luenberger observers and to a class of nonlinear systems using circle criterion based observers. To illustrate the applicability of the supervisory observer, the schemes are applied to a neural mass model to estimate the synaptic gains and the mean membrane potentials of a cortical column.

This work can be extended along two important directions. Firstly, the inclusion of measurement noises and model uncertainties. We expect the proposed supervisory observers to be well-equipped to handle this scenario provided the state-observers of the multi-observer unit are robust in an appropriate sense. Secondly, we mentioned in Section IV that prior information on the localization of the parameters within the parameter set in terms of probability distributions may be used to heuristically select the parameters value in the known parameter set. It would be interesting to assume the existence of such a distribution and to revisit the results of this paper to obtain (stochastic) convergence guarantees which would be potentially stronger than those currently ensured by the available nonlinear Kalman filtering techniques.

APPENDIX

A. Proof of Proposition 1

For any $i \in \{1, \dots, N\}$, $p_i \in \Theta$ and Θ is compact, hence $p_i - p^*$ belongs to some compact set $\tilde{\Theta}$. Let $p_i \in \Theta$ with $i \in \{1, \dots, N\}$ such that $\tilde{p}_i \neq 0$, $\Delta_{\tilde{x}}, \Delta_x, \Delta_u > 0$, $\tilde{x}_i(0) \in \mathcal{H}(0, \Delta_{\tilde{x}})$, $x(0) \in \mathcal{H}(0, \Delta_x)$ and $u \in \mathcal{M}_{\Delta_u}$. We denote the solution to the state error system (5) at time $t \geq 0$ as $\xi(t, \tilde{p}_i)$, where we have omitted its dependence on x, u and $\tilde{x}_i(0)$. Similarly, we denote the output to (5) at time $t \geq 0$ as $\eta(t, \tilde{p}_i)$.

We note that (11) can be expressed in a scalar form as follows for all $t \geq T_f$,

$$\int_{t-T_f}^t (\nu^T \eta(s, \tilde{p}_i))^2 ds \geq \bar{\alpha}_i, \quad (30)$$

where $\nu \in \mathbb{R}^{n_y}$ is any constant vector with $|\nu| = 1$. Since $|\nu^T \eta(s, \tilde{p}_i)| \leq |\nu| |\eta(s, \tilde{p}_i)| = |\eta(s, \tilde{p}_i)|$ and $|\eta(s, \tilde{p}_i)| \leq \sqrt{n_y} |\eta(s, \tilde{p}_i)|_\infty$, inequality (30) implies that

$$n_y \int_{t-T_f}^t |\eta(s, \tilde{p}_i)|_\infty^2 ds \geq \bar{\alpha}_i. \quad (31)$$

Let $W : (t, \tilde{p}_i) \mapsto n_y \int_{t-T_f}^t |\eta(s, \tilde{p}_i)|_\infty^2 ds$ for $t \geq T_f$ and $\tilde{p}_i \in \tilde{\Theta}$. We show that W is continuous in t and \tilde{p}_i . Since f and \hat{f} are continuously differentiable, $\tilde{p}_i \in \tilde{\Theta}$, $p^* \in \Theta$, $u \in \mathcal{M}_{\Delta_u}$ and $x \in \mathcal{M}_{K_x}$ (by Assumption 1), F_i is locally Lipschitz in ξ , uniformly in \tilde{p}_i , p^* , u and x using similar arguments as in the proof of Lemma 2.3 of [24]. Hence, $\xi(t, \tilde{p}_i)$ is continuous in t and \tilde{p}_i by Theorem 2.6 in [24]. As a consequence, we deduce that $\eta(t, \tilde{p}_i)$ is also continuous in t and \tilde{p}_i by using the fact that H is continuous in view of item 2) of Proposition 1. Therefore, we have that $w : (s, \tilde{p}_i) \mapsto n_y |\eta(s, \tilde{p}_i)|_\infty^2$ is continuous in s and \tilde{p}_i . As a consequence, w is uniformly continuous on $[t - T_f, t] \times \tilde{\Theta}$ with $t \geq T_f$. Thus, given $\epsilon > 0$, there exists $\delta = \delta(\epsilon) > 0$ such that for every pair of points $z = (s, \tilde{p}_i)$ and $z' = (s', \tilde{p}'_i)$ such that $|z - z'| < \delta$, $|w(s, \tilde{p}_i) - w(s', \tilde{p}'_i)| < \epsilon$. Let $\epsilon > 0$ and we fix the corresponding constant $\delta > 0$. If $|\tilde{p}_i - \tilde{p}'_i| < \delta$, $\left| \int_{t-T_f}^t w(s, \tilde{p}_i) ds - \int_{t-T_f}^t w(s, \tilde{p}'_i) ds \right| \leq \int_{t-T_f}^t |w(s, \tilde{p}_i) - w(s, \tilde{p}'_i)| ds \leq T_f \epsilon$. Therefore, $W(t, \cdot)$ is continuous on $\tilde{\Theta}$ for any $t \geq T_f$. On the other hand, the continuity of $W(\cdot, \tilde{p}_i)$ for any $\tilde{p}_i \in \tilde{\Theta}$ follows from the continuity of $w(\cdot, \tilde{p}_i)$.

For any $t \geq T_f$ and any $\tilde{p}_i \in \tilde{\Theta} \setminus \{0\}$, $W(t, \tilde{p}_i) \geq \bar{\alpha}_i > 0$ in view of (31). Using the fact that W is continuous, we deduce that there exists a continuous and positive definite function $\tilde{W} : \tilde{\Theta} \rightarrow \mathbb{R}_{\geq 0}$ such that $\tilde{W}(\tilde{p}_i) \leq W(t, \tilde{p}_i)$ for any $t \geq T_f$ and $\tilde{p}_i \in \tilde{\Theta}$. By applying Lemma 3.5 in [24], we derive that there exists $\tilde{\alpha}_{\tilde{y}} \in \mathcal{K}_\infty$ such that $\tilde{W}(\tilde{p}_i) \geq \tilde{\alpha}_{\tilde{y}}(|\tilde{p}_i|)$ for any $\tilde{p}_i \in \tilde{\Theta}$. Hence (10) holds with $\alpha_{\tilde{y}}(s) = n_y^{-1} \tilde{\alpha}_{\tilde{y}}(s)$ (where we use the fact that $|\tilde{p}_i| \geq |\tilde{p}_i|_\infty$ for any \tilde{p}_i). \square

B. Proof of Theorem 1

We first prove that the state error systems (5) and the monitoring signals (8) have desirable properties. Namely, the state error system (5) is shown to satisfy a local ISS property with respect to the parameter error \tilde{p}_i , $i \in \{1, \dots, N\}$, in Lemma 1. This property, along with Assumption 3, are the key ingredients that allow us to show that the monitoring signals are lower and upper

bounded by functions of the parameter error \tilde{p}_i , $i \in \{1, \dots, N\}$, that are strictly increasing with $|\tilde{p}_i|$ in Lemma 2. These lemmas are used to conclude the results of Theorem 1.

Lemma 1. *Consider system (1) and the state error system (4) under Assumptions 1-2. There exist constants $\bar{k}, \bar{\lambda} > 0$ such that for any $\Delta_{\tilde{x}}, \Delta_x, \Delta_u > 0$ there exists a class \mathcal{K}_∞ function $\bar{\gamma}_{\tilde{x}}$ such that for any $p_i \in \Theta$, $i \in \{1, \dots, N\}$, $\tilde{x}_i(0) \in \mathcal{H}(0, \Delta_{\tilde{x}})$, $x(0) \in \mathcal{H}(0, \Delta_x)$, $u \in \mathcal{M}_{\Delta_u}$, the corresponding solution satisfies*

$$|\tilde{x}_i(t)|_\infty \leq \bar{k} \exp(-\bar{\lambda}t) |\tilde{x}_i(0)|_\infty + \bar{\gamma}_{\tilde{x}}(|\tilde{p}_i|_\infty), \quad \forall t \geq 0. \quad (32)$$

□

Proof of Lemma 1: Let $p_i \in \Theta$, where $i \in \{1, \dots, N\}$, $V_i : \mathbb{R}^{n_x} \rightarrow \mathbb{R}_{\geq 0}$, $\tilde{\gamma} : \mathbb{R}^{n_p} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \rightarrow \mathbb{R}_{\geq 0}$, $a_1, a_2, \lambda_0 > 0$ be generated by Assumption 2. Given $\Delta_{\tilde{x}}, \Delta_x, \Delta_u > 0$, let $u \in \mathcal{M}_{\Delta_u}$ and $K_x > 0$ be generated by Assumption 1. Since $\tilde{\gamma}$ is a continuous function and $\tilde{\gamma}(0, x, u) = 0$, for all $x \in \mathbb{R}^{n_x}$, $u \in \mathbb{R}^{n_u}$, $\tilde{\gamma}$ can always be upper bounded by a positive definite function $\tilde{\gamma}(\tilde{p}_i) = \max_{|x|_\infty \leq K_x, |u|_\infty \leq \Delta_u} \tilde{\gamma}(\tilde{p}_i, x, u)$. By Lemma 3.5 in [24] and since the Euclidean and infinity norms are equivalent in \mathbb{R}^{n_p} , there exists a class \mathcal{K}_∞ function $\bar{\gamma}$ such that

$$\tilde{\gamma}_1(\tilde{p}_i) \leq \bar{\gamma}(|\tilde{p}_i|_\infty). \quad (33)$$

We use (6), (7) and (33) to obtain the following for all $|x|_\infty \leq K_x$ and $|u|_\infty \leq \Delta_u$

$$\frac{\partial V_i}{\partial \tilde{x}_i} F_i(\tilde{x}_i, x, \tilde{p}_i, p^*, u) \leq -\lambda_0 V_i(\tilde{x}_i) + \bar{\gamma}(|\tilde{p}_i|_\infty). \quad (34)$$

By the comparison principle (Lemma 2.5 in [24]), for any $\tilde{x}_i(0) \in \mathcal{H}(0, \Delta_{\tilde{x}})$, $x(0) \in \mathcal{H}(0, \Delta_x)$, $u \in \mathcal{M}_{\Delta_u}$, the corresponding solution to (1), (4) verifies

$$V_i(\tilde{x}_i(t)) \leq \exp(-\lambda_0 t) V_i(\tilde{x}_i(0)) + \frac{1}{\lambda_0} \bar{\gamma}(|\tilde{p}_i|_\infty), \quad (35)$$

for $t \geq 0$. We use (6) and the fact that for any $a, b \geq 0$, $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$ to obtain (32) as desired with $\bar{k} := \sqrt{\frac{a_2}{a_1}}$, $\bar{\lambda} := \frac{\lambda_0}{2}$ and $\bar{\gamma}_{\tilde{x}}(r) := \sqrt{\frac{1}{a_1 \lambda_0}} \sqrt{\bar{\gamma}(r)}$ for $r \geq 0$. ■

Lemma 2. *Consider system (1), the state error system (5) and the monitoring signal (8) under Assumptions 1-3. For any $\Delta_{\tilde{x}}, \Delta_x, \Delta_u, \epsilon > 0$, there exist class \mathcal{K}_∞ functions $\underline{\chi}$ and $\bar{\chi}$ independent of ϵ , a constant $T = T(\Delta_{\tilde{x}}, \Delta_x, \Delta_u, \epsilon) > 0$ such that for all $p_i \in \Theta$, $i \in \{1, \dots, N\}$,*

$(x(0), \tilde{x}_i(0)) \in \mathcal{H}(0, \Delta_x) \times \mathcal{H}(0, \Delta_{\tilde{x}})$, for some $u \in \mathcal{M}_{\Delta_u}$ that satisfies Assumption 3, the monitoring signal μ_i in (8) satisfies

$$\underline{\chi}(|\tilde{p}_i|_\infty) \leq \mu_i(t) \leq \bar{\chi}(|\tilde{p}_i|_\infty) + \epsilon, \quad \forall t \geq T. \quad (36)$$

□

Proof of Lemma 2: Let $\Delta_{\tilde{x}}, \Delta_x, \Delta_u, \epsilon > 0$ and $p_i \in \Theta, i \in \{1, \dots, N\}$.

- Since Θ is compact, there exists $K_{\tilde{p}} > 0$ such that for all $p \in \Theta$

$$|p - p^*|_\infty \leq K_{\tilde{p}}, \quad (37)$$

which implies that $|\tilde{p}_i| \leq K_{\tilde{p}}$.

- Let $\bar{k}, \bar{\lambda} > 0$ and the class \mathcal{K}_∞ function $\bar{\gamma}_{\tilde{x}}$ be generated from Lemma 1. By Lemma 1 and (37), the solutions to (1), (5) satisfy the following for all $(\tilde{x}_i(0), x(0)) \in \mathcal{H}(0, \Delta_{\tilde{x}}) \times \mathcal{H}(0, \Delta_x)$, $u \in \mathcal{M}_{\Delta_u}$ and $t \geq 0$,

$$|\tilde{x}_i(t)|_\infty \leq \bar{k} \exp(-\bar{\lambda}t) |\tilde{x}_i(0)|_\infty + \bar{\gamma}_{\tilde{x}}(|\tilde{p}_i|_\infty) \leq \bar{k} \Delta_{\tilde{x}} + \bar{\gamma}_{\tilde{x}}(K_{\tilde{p}}) =: K_{\tilde{x}}. \quad (38)$$

- Since h is continuously differentiable, h is locally Lipschitz. Therefore, there exist constants $l_{\tilde{x}}, l_{\tilde{p}} > 0$ such that for $x \in \mathcal{M}_{K_x}$ and $\tilde{x}_i \in \mathcal{M}_{K_{\tilde{x}}}$ with $K_x, K_{\tilde{x}} > 0$, the following holds for all $t \geq 0$

$$\begin{aligned} |H(\tilde{x}_i(t), x(t), \tilde{p}_i, p^*)|_\infty &= |H(\tilde{x}_i(t), x(t), \tilde{p}_i, p^*) - H(0, x(t), 0, p^*)|_\infty \\ &= |h(\tilde{x}_i(t) + x(t), \tilde{p}_i + p^*) - h(x(t), p^*)|_\infty \\ &\leq |h(\tilde{x}_i(t) + x(t), \tilde{p}_i + p^*) - h(x(t), p^*)| \\ &\leq l_{\tilde{x}} |\tilde{x}_i(t)| + l_{\tilde{p}} |\tilde{p}_i| \\ &\leq \sqrt{n_x} l_{\tilde{x}} |\tilde{x}_i(t)|_\infty + \sqrt{n_p} l_{\tilde{p}} |\tilde{p}_i|_\infty, \end{aligned} \quad (39)$$

where we have used that $H(0, x(t), 0, p^*) = 0$ in view of the definition of H in (5).

- Let $\lambda > 0$ come from (8), we choose $\epsilon_{\tilde{x}}, \epsilon_\mu > 0$ sufficiently small such that

$$\epsilon_\mu + \frac{4n_x l_{\tilde{x}}^2 \epsilon_{\tilde{x}}^2}{\lambda} = \epsilon. \quad (40)$$

Let $T_{\tilde{x}} > 0$ be sufficiently large such that

$$\bar{k} \exp(-\bar{\lambda}t) \Delta_{\tilde{x}} \leq \epsilon_{\tilde{x}}, \quad \forall t \geq T_{\tilde{x}}. \quad (41)$$

Let $T_\mu \geq T_{\tilde{x}}$ be sufficiently large such that

$$\frac{1}{\lambda} \exp(-\lambda(T_\mu - T_{\tilde{x}}))(\sqrt{n_x}l_{\tilde{x}}K_{\tilde{x}} + \sqrt{n_p}l_{\tilde{p}}K_{\tilde{p}})^2 \leq \epsilon_\mu. \quad (42)$$

- Let $T_f > 0$ and the class \mathcal{K}_∞ function $\alpha_{\tilde{y}}$ be generated by Assumption 3 such that for all $\tilde{x}_i(0) \in \mathcal{H}(0, \Delta_{\tilde{x}})$, for some $u \in \mathcal{M}_{\Delta_u}$, for all $x \in \mathcal{M}_{K_x}$, the solution to (5) satisfies

$$\int_{t-T_f}^t |H(\tilde{x}_i(s), x(s), \tilde{p}_i, p^*)|_\infty^2 ds \geq \alpha_{\tilde{y}}(|\tilde{p}_i|_\infty), \quad \forall t \geq T_f. \quad (43)$$

- We define the following

$$T := \max\{T_\mu, T_{\tilde{x}}, T_f\}, \quad (44)$$

$$\underline{\chi}(r) := \exp(-\lambda T_f) \alpha_{\tilde{y}}(r) \quad \forall r \geq 0 \quad (45)$$

$$\bar{\chi}(r) := \frac{4n_x l_{\tilde{x}}^2}{\lambda} \bar{\gamma}_{\tilde{x}}^2(r) + \frac{2n_p l_{\tilde{p}}^2}{\lambda} r^2 \quad \forall r \geq 0. \quad (46)$$

Note that the class \mathcal{K}_∞ functions $\underline{\chi}$ and $\bar{\chi}$ depend only on $\Delta_{\tilde{x}}$, Δ_x , Δ_u and not on ϵ . Let $t \geq T$, where T is defined in (44). By definition of the monitoring signals in (8),

$$\mu_i(t) = \int_0^t \exp(-\lambda(t-s)) |H(\tilde{x}_i(s), x(s), \tilde{p}_i, p^*)|_\infty^2 ds. \quad (47)$$

We first establish the desired lower bound on μ_i

$$\begin{aligned} \mu_i(t) &= \int_0^{t-T_f} \exp(-\lambda(t-s)) |H(\tilde{x}_i(s), x(s), \tilde{p}_i, p^*)|_\infty^2 ds \\ &\quad + \int_{t-T_f}^t \exp(-\lambda(t-s)) |H(\tilde{x}_i(s), x(s), \tilde{p}_i, p^*)|_\infty^2 ds \end{aligned} \quad (48)$$

$$\geq \int_{t-T_f}^t \exp(-\lambda(t-s)) |H(\tilde{x}_i(s), x(s), \tilde{p}_i, p^*)|_\infty^2 ds. \quad (49)$$

As $s \mapsto \exp(\lambda s)$ is strictly increasing,

$$\mu_i(t) \geq \exp(-\lambda T_f) \int_{t-T_f}^t |H(\tilde{x}_i(s), x(s), \tilde{p}_i, p^*)|_\infty^2 ds. \quad (50)$$

From (43), (45), since $t \geq T_f$,

$$\mu_i(t) \geq \exp(-\lambda T_f) \alpha_{\tilde{y}}(|\tilde{p}_i|_\infty) = \underline{\chi}(|\tilde{p}_i|_\infty). \quad (51)$$

We now obtain the desired upper bound of μ_i

$$\begin{aligned} \mu_i(t) &= \int_0^{T_{\tilde{x}}} \exp(-\lambda(t-s)) |H(\tilde{x}_i(s), x(s), \tilde{p}_i, p^*)|_\infty^2 ds \\ &\quad + \int_{T_{\tilde{x}}}^t \exp(-\lambda(t-s)) |H(\tilde{x}_i(s), x(s), \tilde{p}_i, p^*)|_\infty^2 ds. \end{aligned} \quad (52)$$

Using (39) and the fact that for any $a, b \geq 0$, $(a + b)^2 \leq 2a^2 + 2b^2$, we obtain

$$\begin{aligned} \mu_i(t) &\leq \frac{1}{\lambda} \exp(-\lambda(t - T_{\bar{x}})) \left(\sup_{s \in [0, T_{\bar{x}}]} |H(\tilde{x}_i(s), x(s), \tilde{p}_i, p^*)|_\infty^2 \right) \\ &\quad + \int_{T_{\bar{x}}}^t \exp(-\lambda(t - s)) \left(\sqrt{n_x} l_{\bar{x}} |\tilde{x}_i(s)|_\infty + \sqrt{n_p} l_{\bar{p}} |\tilde{p}_i|_\infty \right)^2 ds \\ &\leq \frac{1}{\lambda} \exp(-\lambda(t - T_{\bar{x}})) \left(\sup_{s \in [0, T_{\bar{x}}]} |H(\tilde{x}_i(s), x(s), \tilde{p}_i, p^*)|_\infty^2 \right) \\ &\quad + \int_{T_{\bar{x}}}^t \exp(-\lambda(t - s)) \left(2n_x l_{\bar{x}}^2 |\tilde{x}_i(s)|_\infty^2 + 2n_p l_{\bar{p}}^2 |\tilde{p}_i|_\infty^2 \right) ds. \end{aligned} \quad (53)$$

We also have from Lemma 1 and (41) that, as $t \geq T_{\bar{x}}$,

$$|\tilde{x}_i(t)|_\infty \leq \epsilon_{\bar{x}} + \gamma_{\bar{x}}(|\tilde{p}_i|_\infty), \quad (54)$$

which implies that, using the fact that for any $a, b > 0$, $(a + b)^2 \leq 2a^2 + 2b^2$

$$|\tilde{x}_i(t)|_\infty^2 \leq 2\epsilon_{\bar{x}}^2 + 2\gamma_{\bar{x}}^2(|\tilde{p}_i|_\infty). \quad (55)$$

Also, by (38) and (39), we obtain

$$\begin{aligned} \sup_{s \in [0, T_{\bar{x}}]} |H(\tilde{x}_i(s), x(s), \tilde{p}_i, p^*)|_\infty^2 &\leq \sup_{s \in [0, T_{\bar{x}}]} \left(\sqrt{n_x} l_{\bar{x}} |\tilde{x}_i(s)|_\infty + \sqrt{n_p} l_{\bar{p}} |\tilde{p}_i|_\infty \right)^2 \\ &= (\sqrt{n_x} l_{\bar{x}} K_{\bar{x}} + \sqrt{n_p} l_{\bar{p}} K_{\bar{p}})^2. \end{aligned} \quad (56)$$

Hence, from (53) and in view of (42), (55) and (56),

$$\begin{aligned} \mu_i(t) &\leq \epsilon_\mu + \int_{T_{\bar{x}}}^t \exp(-\lambda(t - s)) \left(4n_x l_{\bar{x}}^2 \epsilon_{\bar{x}}^2 + 4n_x l_{\bar{x}}^2 \gamma_{\bar{x}}^2(|\tilde{p}_i|_\infty) + 2n_p l_{\bar{p}}^2 |\tilde{p}_i|_\infty^2 \right) ds \\ &= \epsilon_\mu + \left(\int_{T_{\bar{x}}}^t \exp(-\lambda(t - s)) ds \right) \left(4n_x l_{\bar{x}}^2 \epsilon_{\bar{x}}^2 + 4n_x l_{\bar{x}}^2 \gamma_{\bar{x}}^2(|\tilde{p}_i|_\infty) + 2n_p l_{\bar{p}}^2 |\tilde{p}_i|_\infty^2 \right). \end{aligned} \quad (57)$$

As $\int_{T_{\bar{x}}}^t \exp(-\lambda(t - s)) ds = \frac{1}{\lambda} (1 - \exp(-\lambda(t - T_{\bar{x}}))) \leq \frac{1}{\lambda}$. Therefore, from (57),

$$\mu_i(t) \leq \epsilon_\mu + \frac{4n_x l_{\bar{x}}^2 \epsilon_{\bar{x}}^2}{\lambda} + \frac{4n_x l_{\bar{x}}^2}{\lambda} \gamma_{\bar{x}}^2(|\tilde{p}_i|_\infty) + \frac{2n_p l_{\bar{p}}^2}{\lambda} |\tilde{p}_i|_\infty^2. \quad (58)$$

By the definition of ϵ and $\bar{\chi}$ in (40) and (46) respectively

$$\mu_i(t) \leq \epsilon + \bar{\chi}(|\tilde{p}_i|_\infty). \quad (59)$$

Therefore, (36) holds in view of (51) and (59). ■

Proof of Theorem 1: Given $\Delta_{\bar{x}}, \Delta_x, \Delta_u > 0$ and $\nu_{\bar{p}}, \nu_{\bar{x}} > 0$, we construct the ingredients needed.

- Let $\nu_1 > 0$ be sufficiently small such that

$$\nu_1 \leq \bar{\gamma}_{\tilde{x}}^{-1}(\nu_{\tilde{x}}), \quad (60)$$

where $\bar{\gamma}_{\tilde{x}}$ is the class \mathcal{K}_∞ function generated by Lemma 1.

- Let the class \mathcal{K}_∞ functions $\bar{\chi}$ and $\underline{\chi}$ be generated by Lemma 2. Choose $\epsilon \in (0, \underline{\chi}(\min\{\nu_1, \nu_{\tilde{p}}\})]$ and

$$d^* := \bar{\chi}^{-1}(\underline{\chi}(\min\{\nu_1, \nu_{\tilde{p}}\}) - \epsilon). \quad (61)$$

Recall that the chosen monitoring signal at any time $t \geq 0$ is denoted as $\mu_{\sigma(t)}(t)$. By the definition of the selection criterion (12),

$$\mu_{\sigma(t)}(t) \leq \mu_j(t), \quad \forall t \geq 0, \forall j \in \{1, \dots, N\}. \quad (62)$$

Since (62) holds, we consider the monitoring signal with the smallest parameter estimation error, i.e. $j^* = \arg \min_{j \in \{1, \dots, N\}} |\tilde{p}_j|_\infty$ and $|\tilde{p}_{j^*}| = d(p^*, \hat{\Theta})$ where $d(p^*, \hat{\Theta})$ is defined in (3). Using Lemma 2, there exists $T > 0$ such that

$$\underline{\chi}(|\tilde{p}_{\sigma(t)}(t)|_\infty) \leq \mu_{\sigma(t)}(t) \leq \mu_{j^*}(t) \leq \bar{\chi}(|\tilde{p}_{j^*}|_\infty) + \epsilon, \quad \forall t \geq T. \quad (63)$$

Therefore

$$|\tilde{p}_{\sigma(t)}(t)|_\infty \leq \underline{\chi}^{-1}(\bar{\chi}(d(p^*, \hat{\Theta})) + \epsilon), \quad \forall t \geq T. \quad (64)$$

Recall from Section IV-A1 that we sample the parameter set Θ in a manner such that $d(p^*, \hat{\Theta}) \rightarrow 0$ as $N \rightarrow \infty$. Therefore, there is an $N^* \in \mathbb{N}_{\geq 1}$ such that $d(p^*, \hat{\Theta}) \leq d^*$ for all $N \geq N^*$. We fix N^* and take $N \geq N^*$. Hence, for all $t \geq T$

$$|\tilde{p}_{\sigma(t)}(t)|_\infty \leq \underline{\chi}^{-1}(\bar{\chi}(d^*) + \epsilon) = \min\{\nu_1, \nu_{\tilde{p}}\} \leq \nu_{\tilde{p}}. \quad (65)$$

We now examine the state estimation error $\tilde{x}_{\sigma(t)}(t)$. For all $t \geq 0$,

$$|\tilde{x}_{\sigma(t)}(t)|_\infty \leq \max_{i \in \{1, \dots, N\}} |\tilde{x}_i(t)|_\infty. \quad (66)$$

By Lemma 1, we have that for $i \in \{1, \dots, N\}$, for all $\tilde{x}_i(0) \in \mathcal{H}(0, \Delta_{\tilde{x}})$, $x(0) \in \mathcal{H}(0, \Delta_x)$ and $u \in \mathcal{M}_{\Delta_u}$, the solution to (1), (5) satisfies

$$|\tilde{x}_i(t)|_\infty \leq \bar{k} \exp(-\bar{\lambda}t) |\tilde{x}_i(0)|_\infty + \bar{\gamma}_{\tilde{x}}(|\tilde{p}_i|_\infty), \quad (67)$$

where $\bar{k}, \bar{\lambda} > 0$ and $\bar{\gamma}_{\tilde{x}} \in \mathcal{K}_\infty$. Like in the proof of Lemma 2, since $p^*, p_i \in \Theta$, there exists $K_{\tilde{p}} > 0$, which is independent of i , such that $|\tilde{p}_i|_\infty \leq K_{\tilde{p}}$. Hence, from (66) and (67), we obtain for all $t \geq 0$

$$\begin{aligned} |\tilde{x}_i(t)|_\infty &\leq \bar{k}\Delta_{\tilde{x}} + \bar{\gamma}_{\tilde{x}} \left(\max_{i \in \{1, \dots, N\}} |\tilde{p}_i|_\infty \right) \\ &= \bar{k}\Delta_{\tilde{x}} + \bar{\gamma}_{\tilde{x}}(K_{\tilde{p}}) =: \bar{K}_{\tilde{x}}. \end{aligned} \quad (68)$$

Hence, we have from (66) and (68) that the solution to the chosen state estimation error system satisfies

$$|\tilde{x}_{\sigma(t)}(t)|_\infty \leq \bar{K}_{\tilde{x}}, \quad \forall t \geq 0. \quad (69)$$

Furthermore, we have from (65) that the parameter error \tilde{p}_σ converges to the hypercube centered at 0 and of edge length $2\nu_{\tilde{p}}$ in finite-time. Let $\mathcal{S} := \{i \in \{1, \dots, N\} : |\tilde{p}_i| \leq \min\{\nu_{\tilde{p}}, \nu_1\}\}$. Therefore, as $t \geq T$

$$|\tilde{x}_{\sigma(t)}(t)|_\infty \leq \max_{i \in \mathcal{S}} |\tilde{x}_i(t)|_\infty. \quad (70)$$

Consequently, we have from (67) that

$$\limsup_{t \rightarrow \infty} |\tilde{x}_{\sigma(t)}(t)|_\infty \leq \limsup_{t \rightarrow \infty} \max_{i \in \mathcal{S}} |\tilde{x}_i(t)|_\infty \leq \max_{i \in \mathcal{S}} \bar{\gamma}_{\tilde{x}}(|\tilde{p}_i|_\infty) \leq \bar{\gamma}_{\tilde{x}}(\min\{\nu_{\tilde{p}}, \nu_1\}) \leq \bar{\gamma}_{\tilde{x}}(\nu_1).$$

By (60), we obtain

$$\limsup_{t \rightarrow \infty} |\tilde{x}_{\sigma(t)}(t)|_\infty \leq \nu_{\tilde{x}}. \quad (71)$$

Finally, we have shown (15) in view of (65), (69) and (71). ■

C. Proof of Theorem 2

Let $\Delta_{\tilde{x}}, \Delta_x, \Delta_u, \nu_{\tilde{x}}, \nu_{\tilde{p}} > 0$ and $\alpha \in (0, 1)$. The solutions to (1) and (21) are unique and are defined for all positive time, for any initial condition and any input u in view of Sections III and IV-A2. On the other hand, $p_i(t) \in \Theta$ for any $i \in \{1, \dots, N\}$ and $t \geq 0$ in view of Section V-A. Hence, we will be able to invoke the results of Lemmas 1-2 to derive the desired result.

Let $\nu_1 > 0$ be sufficiently small such that

$$\nu_1 \leq \min\{\nu_{\tilde{p}}, \bar{\gamma}_{\tilde{x}}^{-1}(\nu_{\tilde{x}})\}, \quad (72)$$

where $\bar{\gamma}_{\tilde{x}}$ is a class \mathcal{K}_∞ function generated by Lemma 1. Let $\underline{\chi}$ and $\bar{\chi}$ be the class \mathcal{K}_∞ functions generated by Lemma 2 (which are independent of the choice of the parameters $p_i, i \in \{1, \dots, N\}$, see Lemma 2). We introduce $\Delta_\infty \in (0, \Delta)$ which is sufficiently small such that

$$\underline{\chi}^{-1} \circ \bar{\chi}(\pi(\Delta_\infty, 0)) < \nu_1 \quad \text{and} \quad \Delta_\infty < \nu_1, \quad (73)$$

where π is the class \mathcal{KL} function in (17). It is always possible to ensure (73) as $\underline{\chi}, \bar{\chi} \in \mathcal{K}_\infty$ and $\pi \in \mathcal{KL}$. We select $T > 0$ sufficiently large such that the conclusion of Lemma 2 holds with $\epsilon > 0$ sufficiently small such that

$$\underline{\chi}^{-1} \circ \bar{\chi}(\pi(\Delta_\infty, 0) + \epsilon) < \nu_1 \quad \text{and} \quad \underline{\chi}^{-1}(\epsilon) < \alpha\Delta_\infty < \nu_1. \quad (74)$$

We take $N^* \in \mathbb{N}_{\geq 1}$ sufficiently large such that

$$\underline{\chi}^{-1}(\bar{\chi}(\pi(s, N^*)) + \epsilon) \leq \alpha s \quad \forall s \in [\Delta_\infty, \Delta], \quad (75)$$

which is always possible as $\underline{\chi}, \bar{\chi} \in \mathcal{K}_\infty$, $\pi \in \mathcal{KL}$ and $\underline{\chi}^{-1}(\epsilon) < \alpha\Delta_\infty \leq \alpha s$ for $s \in [\Delta_\infty, \Delta]$ according to (74). A sufficient condition to ensure (75) is $\underline{\chi}(\bar{\chi}(\pi(\Delta, N^*)) + \epsilon) \leq \alpha\Delta_\infty$ for example.

Let $N \geq N^*$, $T^* = T$, $T_d \geq T^*$, $(x(0), \tilde{x}_i(0)) \in \mathcal{H}(0, \Delta_x) \times \mathcal{H}(0, \Delta_{\tilde{x}})$ for $i \in \{1, \dots, N\}$ and $u \in \mathcal{M}_{\Delta_u}$ such that Assumption 3 holds. We proceed by induction. Like in (64), using Lemma 2, we derive that, since $t_1 = T_d \geq T$ in view of (16) and since $p^* \in \Theta(0) = \Theta$,

$$\begin{aligned} |\tilde{p}_{\sigma(t_1^-)}(t_1^-)|_\infty &\leq \underline{\chi}^{-1}(\bar{\chi}(d(p^*, \hat{\Theta}(0)) + \epsilon)) \\ &\leq \underline{\chi}^{-1}(\bar{\chi}(\pi(\Delta(0), N) + \epsilon)). \end{aligned} \quad (76)$$

We know that $\Delta_\infty < \Delta(0) = \Delta$, hence, in view of (75)

$$|\tilde{p}_{\sigma(t_1^-)}(t_1^-)|_\infty \leq \alpha\Delta(0) = \Delta(1). \quad (77)$$

We deduce from (77) that $p^* \in \mathcal{H}(\hat{p}(t_1^-), \Delta(1))$. On the other hand, $p^* \in \Theta(0)$ which implies that $p^* \in \Theta(1) = \mathcal{H}(\hat{p}(t_1^-), \Delta(1)) \cap \Theta(0)$.

Let $k \in \mathbb{N}_{\geq 1}$. Since $t_{k+1} - t_k = T_d \geq T$ in view of (16), we obtain as in (76)

$$\begin{aligned} |\tilde{p}_{\sigma(t_k^-)}(t_k^-)|_\infty &\leq \underline{\chi}^{-1}(\bar{\chi}(d(p^*, \hat{\Theta}(k-1)) + \epsilon)) \\ &\leq \underline{\chi}^{-1}(\bar{\chi}(\pi(\Delta(k-1), N) + \epsilon)). \end{aligned} \quad (78)$$

If $\Delta(k-1) \leq \Delta_\infty$, then, in view of (73),

$$|\tilde{p}_{\sigma(t_k^-)}(t_k^-)|_\infty \leq \Delta_\infty \leq \nu_1 \leq \nu_{\tilde{p}}. \quad (79)$$

If $\Delta(k-1) > \Delta_\infty$, then, in view of (75)

$$|\tilde{p}_{\sigma(t_k^-)}(t_k^-)|_\infty \leq \alpha\Delta(k-1) = \Delta(k). \quad (80)$$

In this case $p^* \in \mathcal{H}(\hat{p}(t_k^-), \Delta(k))$ and we know that $p^* \in \Theta(0) \cap \dots \cap \Theta(k-1)$. As a consequence, $p^* \in \Theta(k)$.

We have shown that for any $k \in \mathbb{N}_{\geq 1}$

$$|\tilde{p}_{\sigma(t_k^-)}(t_k^-)|_{\infty} \leq \begin{cases} \nu_1 & \text{when } \Delta(k-1) \leq \Delta_{\infty} \\ \Delta(k) & \text{when } \Delta(k-1) > \Delta_{\infty}. \end{cases} \quad (81)$$

Since $k \rightarrow \infty$ in view of (16) and $\Delta(k) = \alpha^k \Delta \rightarrow 0$ as $k \rightarrow \infty$, we deduce that there exists $\bar{T} > 0$ such that

$$|\tilde{p}_{\sigma(t)}(t)| \leq \nu_1 \leq \nu_{\tilde{p}} \quad \forall t \geq \bar{T}. \quad (82)$$

The proof is completed by following the same lines as in the proof of Theorem 1. \blacksquare

D. Proof of Proposition 2

For any $p_i \in \Theta$, $i \in \{1, \dots, N\}$, we obtain the following state estimation error system

$$\begin{aligned} \dot{\tilde{x}}_i &= \left(A(p_i) + L(p_i)C(p_i) \right) \hat{x}_i - \left(A(p^*) + L(p_i)C(p^*) \right) x + \left(B(p_i) - B(p^*) \right) u \\ &= \left(A(p_i) + L(p_i)C(p_i) \right) \tilde{x}_i + \left(\tilde{A}(p_i, p^*) + L(p_i)\tilde{C}(p_i, p^*) \right) x + \tilde{B}(p_i, p^*)u, \end{aligned} \quad (83)$$

where we denote $\tilde{A}(p_i, p^*) := A(p_i) - A(p^*)$, $\tilde{B}(p_i, p^*) := B(p_i) - B(p^*)$ and $\tilde{C}(p_i, p^*) := C(p_i) - C(p^*)$. Let $p_i \in \Theta$ and $V_i : \tilde{x}_i \mapsto \tilde{x}_i^T P_i \tilde{x}_i$, where P_i is a real symmetric, positive definite matrix which satisfies

$$P_i \left(A(p_i) + L(p_i)C(p_i) \right) + \left(A(p_i) + L(p_i)C(p_i) \right)^T P_i = -\nu_i \mathbb{I}, \text{ with } \nu_i > 1. \quad (84)$$

Such a matrix P_i always exists according to Theorem 3.6 of [24] since $A(p_i) + L(p_i)C(p_i)$ is Hurwitz. Hence, (6) is satisfied with $a_1 := \min_{i \in \{1, \dots, N\}} \lambda_{\min}(P_i)$ and $a_2 := n_x \max_{i \in \{1, \dots, N\}} \lambda_{\max}(P_i)$.

Let $\tilde{x}_i, x \in \mathbb{R}^{n_x}$ and $u \in \mathbb{R}^{n_u}$, it holds that

$$\begin{aligned} \langle \nabla V_i(\tilde{x}_i), F_i(\tilde{x}_i, x, \tilde{p}_i, p^*, u) \rangle &= \tilde{x}_i^T \left(P_i \left(A(p_i) + L(p_i)C(p_i) \right) + \left(A(p_i) + L(p_i)C(p_i) \right)^T P_i \right) \tilde{x}_i \\ &\quad + 2\tilde{x}_i^T P_i \left(\left(\tilde{A}(\tilde{p}_i + p^*, p^*) + L(p_i)\tilde{C}(\tilde{p}_i + p^*, p^*) \right) x + \tilde{B}(\tilde{p}_i + p^*, p^*)u \right) \\ &\leq -\nu_i |\tilde{x}_i|^2 + 2 |\tilde{x}_i| |P_i| \gamma(\tilde{p}_i, x, u), \end{aligned} \quad (85)$$

where $\gamma(\tilde{p}_i, x, u) := \max_{p \in \Theta} \left| \left(\tilde{A}(\tilde{p}_i + p, p) + L(p_i)\tilde{C}(\tilde{p}_i + p, p) \right) x + \tilde{B}(\tilde{p}_i + p, p^*)u \right|$. The function γ is continuous since A , B and C are continuous in their argument and Θ is compact set. Moreover $\gamma(0, x, u) = 0$ for any $x \in \mathbb{R}^{n_x}$ and $u \in \mathbb{R}^{n_u}$. Using the fact that $2ab \leq \frac{\nu_i}{2}a^2 + \frac{2}{\nu_i}b^2$ for any $a, b \in \mathbb{R}$, we deduce from (85) that, for any $\tilde{x}_i, x \in \mathbb{R}^{n_x}$ and $u \in \mathbb{R}^{n_u}$,

$$\langle \nabla V_i(\tilde{x}_i), F_i(\tilde{x}_i, x, \tilde{p}_i, p^*, u) \rangle \leq -\frac{\nu_i}{2} |\tilde{x}_i|^2 + \frac{2}{\nu_i} |P_i|^2 \gamma(\tilde{p}_i, x, u)^2 \quad (86)$$

from which we derive that (7) is satisfied by using (6) and by invoking the equivalence of the infinity and the Euclidean norms. \square

E. Proof of Proposition 4

Let $p_i \in \Theta$ with $i \in \{1, \dots, N\}$, we obtain the following state estimation error system from (26) and (28)

$$\begin{aligned} \dot{\tilde{x}}_i = & \left(A(p_i) + L(p_i)C(p_i) \right) \tilde{x}_i + G(p_i) \left(\gamma(w_i) - \gamma(v) \right) \\ & + \tilde{G}(p_i, p^*) \gamma(v) + \left(\tilde{A}(p_i, p^*) + L(p_i) \tilde{C}(p_i, p^*) \right) x + \tilde{B}(p_i, p^*) \phi(u, y), \end{aligned} \quad (87)$$

where $v := Hx$, $w_i := H\hat{x}_i + K(p_i)(C(p_i)\hat{x}_i - y)$, $\tilde{A}(p_i, p^*) := A(p_i) - A(p^*)$, $\tilde{B}(p_i, p^*) := B(p_i) - B(p^*)$, $\tilde{G}(p_i, p^*) := G(p_i) - G(p^*)$ and $\tilde{C}(p_i, p^*) := C(p_i) - C(p^*)$. In view of (4) and according to the mean value theorem, there exists $\delta(t) = \text{diag}(\delta_1(t), \dots, \delta_{n_\gamma}(t))$, where $\delta_k(t)$ take values in the interval $[a_{\gamma_k}, b_{\gamma_k}]$ so that, for $\gamma = (\gamma_1, \dots, \gamma_{n_\gamma})$,

$$\gamma(w_i) - \gamma(v) = \delta(t)(w_i - v), \quad \forall w_i, v \in \mathbb{R}^{n_\gamma}, i \in \{1, \dots, N\}. \quad (88)$$

We define $V_i : \tilde{x}_i \mapsto \tilde{x}_i^T P_i \tilde{x}_i$, where P_i is a real symmetric, positive definite matrix given by (29). Note that V_i satisfies inequality (6) of Assumption 2 with $a_1 = \lambda_{\min}(P_i)$ and $a_2 = n_x \lambda_{\max}(P_i)$. By following the proof of Theorem 2 in [10] with the vector $\chi_i := (\tilde{x}_i, \delta(t)(H + K(p_i)C)\tilde{x}_i, \bar{w})$, where $\bar{w} = \bar{w}(p_i, p^*, x, u) := \tilde{G}(p_i, p^*)\gamma(Hx) + (\tilde{A}(p_i, p^*) + L(p_i)\tilde{C}(p_i, p^*))x + \tilde{B}(p_i, p^*)\phi(u, y)$, we obtain, for any $\tilde{x}_i, x \in \mathbb{R}^{n_x}$, $u \in \mathbb{R}^{n_u}$,

$$\langle \nabla V_i(\tilde{x}_i), F_i(\tilde{x}_i, x, \tilde{p}_i, p^*, u) \rangle \leq -\nu_i |\tilde{x}_i|^2 + \mu_i |\bar{w}|^2. \quad (89)$$

We then use the same arguments as in the proof of Proposition 2 to derive (7). \square

REFERENCES

- [1] V. Adetola and M. Guay. Finite-time parameter estimation in adaptive control of nonlinear systems. *IEEE Transactions on Automatic Control*, 53(3):807–811, 2008.
- [2] A.P. Aguiar, M. Athans, and A. Pascoal. Convergence properties of a continuous-time multiple-model adaptive estimator. In *Proc. of European Control Conference*, 2007.
- [3] A.P. Aguiar, V. Hassani, A.M. Pascoal, and M. Athans. Identification and convergence analysis of a class of continuous-time multiple-model adaptive estimators. In *Proc. of The 17th IFAC World Congress*, 2008.
- [4] B.D.O. Anderson and J.B. Moore. *Optimal filtering*. Prentice-Hall Englewood Cliffs, NJ, 1979.

- [5] V. Andrieu and L. Praly. On the Existence of a Kazantzis–Kravaris/Luenberger Observer. *SIAM Journal on Control and Optimization*, 45(2):432–456, 2006.
- [6] M. Arcak and P. Kokotović. Nonlinear observers: a circle criterion design and robustness analysis. *Automatica*, 37(12):1923–1930, 2001.
- [7] M. Arcak and P. Kokotović. Observer-based control of systems with slope-restricted nonlinearities. *IEEE Transactions on Automatic Control*, 46:1, 2001.
- [8] G. Battistelli, J.P. Hespanha, and P. Tesi. Supervisory control of switched nonlinear systems. *International Journal of Adaptive Control and Signal Processing*, 2012.
- [9] G. Besançon. *Nonlinear observers and applications*. Lecture notes in control and information sciences. Springer, 2007.
- [10] M.S. Chong, R. Postoyan, D. Nešić, L. Kuhlmann, and A. Varsavsky. A robust circle criterion observer with application to neural mass models. *Automatica*, 48(11):2986–2989, 2012.
- [11] M. Farza, M. M’Saad, T. Maatoug, and M. Kamoun. Adaptive observers for nonlinearly parameterized class of nonlinear systems. *Automatica*, 45(10):2292–2299, 2009.
- [12] R. Goebel, R.G. Sanfelice, and A. Teel. Hybrid dynamical systems. *IEEE Control Systems*, 29(2):28–93, 2009.
- [13] L. Grüne, E.D. Sontag, and F.R. Wirth. *Systems & Control Letters*, 38(2):127–134, 1999.
- [14] Z. Han and K.S. Narendra. New concepts in adaptive control using multiple models. *IEEE Transactions on Automatic Control*, 57(1):78–89, 2012.
- [15] M. Hartman, N. Bauer, and A. Teel. Robust finite-time parameter estimation using a hybrid systems framework. *IEEE Transactions on Automatic Control*, 57:2956–2962, 2012.
- [16] J.P. Hespanha, D. Liberzon, and A. S. Morse. Hysteresis-based switching algorithms for supervisory control of uncertain systems. *Automatica*, 39(2):263–272, February 2003.
- [17] J.P. Hespanha, D. Liberzon, and A.S. Morse. Overcoming the limitations of adaptive control by means of logic-based switching. *Systems & Control Letters*, 49(1):49–65, 2003.
- [18] J.P. Hespanha, D. Liberzon, A.S. Morse, B.D.O. Anderson, T.S. Brinsmead, and F. De Bruyne. Multiple model adaptive control. part 2: switching. *International journal of robust and nonlinear control*, 11(5):479–496, 2001.
- [19] P.A. Ioannou and J. Sun. *Robust adaptive control*. Prentice Hall, 1996.
- [20] B.H. Jansen and V.G. Rit. Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biological Cybernetics*, 73:357–366, 1995.
- [21] S.J. Julier and J.K. Uhlmann. Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, 92(3):401–422, 2004.
- [22] I. Karafyllis and Z.P. Jiang. Hybrid dead-beat observers for a class of nonlinear systems. *Systems & Control Letters*, 60(8):608–617, 2011.
- [23] D. Karagiannis, D. Carnevale, and A. Astolfi. Invariant manifold based reduced-order observer design for nonlinear systems. *IEEE Transactions on Automatic Control*, 53(11):2602–2614, 2008.
- [24] H.K. Khalil. *Nonlinear systems*. Prentice Hall, 2nd edition, 1996.
- [25] H.K. Khalil. High-gain observers in nonlinear feedback control. *New directions in nonlinear observer design*, pages 249–268, 1999.
- [26] X. Li and Y. Bar-Shalom. Multiple-model estimation with variable structure. *IEEE Transactions on Automatic Control*, 41(4):478–493, 1996.
- [27] D. Liberzon. Hybrid feedback stabilization of systems with quantized signals. *Automatica*, 39(9):1543–1554, 2003.

- [28] D. Liberzon. *Switching in systems and control*. Springer, 2003.
- [29] L. Ljung. *System identification*. Wiley Online Library, 1999.
- [30] A.S. Morse. Supervisory control of families of linear set-point controllers. 2. robustness. *IEEE Transactions on Automatic Control*, 42(11):1500–1515, 1997.
- [31] K.S. Narendra and A.M. Annaswamy. *Stable adaptive systems*. Prentice-Hall, 1989.
- [32] D. Nešić, L. Zaccarian, and A.R. Teel. Stability properties of reset systems. *Automatica*, 44(8):2019–2026, 2008.
- [33] E. Panteley, A. Loria, and A. Teel. Relaxed persistency of excitation for uniform asymptotic stability. *IEEE Transactions on Automatic Control*, 46(12):1874–1886, 2001.
- [34] C. Prieur. Uniting local and global controllers with robustness to vanishing noise. *Mathematics of Control, Signals, and Systems (MCSS)*, 14(2):143–172, 2001.
- [35] S. Sastry and M. Bodson. *Adaptive control: stability, convergence and robustness*. Dover Publications, 2011.
- [36] S.J. Schiff. *Neural Control Engineering: The Emerging Intersection Between Control Theory and Neuroscience*. Computational Neuroscience. The MIT Press, 2011.
- [37] I.Y. Tyukin, D.V. Prokhorov, and C. van Leeuwen. Adaptation and parameter estimation in systems with unstable target dynamics and nonlinear parametrization. *IEEE Transactions on Automatic Control*, 52(9):1543–1559, 2007.
- [38] L. Vu, D. Chatterjee, and D. Liberzon. Input-to-state stability of switched systems and switching adaptive control. *Automatica*, 43:649–646, 2007.
- [39] L. Vu and D. Liberzon. Supervisory control of uncertain linear time-varying systems. *IEEE Transactions on Automatic Control*, 56(1):27–42, 2011.
- [40] F. Wendling, A. Hernandez, J.J. Bellanger, P. Chauvel, and F. Bartolomei. Interictal to ictal transition in human temporal lobe epilepsy: insights from a computational model of intracerebral EEG. *Journal of Clinical Neurophysiology*, 22(5):343, 2005.
- [41] A. Zemouche and M. Boutayeb. A unified H_∞ adaptive observer synthesis method for a class of systems with both Lipschitz and monotone nonlinearities. *Systems & Control Letters*, 58(4):282–288, 2009.
- [42] Q.H. Zhang. Adaptive observer for multiple-input-multiple-output (MIMO) linear time-varying systems. *IEEE Transactions on Automatic Control*, 47(3):525–529, 2002.